Characterization of the Radiation Desiccation Response Regulon of the Radioresistant Bacterium *Deinococcus radiodurans* by Integrative Genomic Analyses

Nicolas Eugénie [ID], Yvan Zivanovic [ID], Gaelle Lelandais, Geneviève Coste, Claire Bouthier de la Tour, Esma Bentchikou, Pascale Servant † and Fabrice Confalonieri *,†

Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France; nicolas.eugenie@i2bc.paris-saclay.fr (N.E.); yvan.zivanovic@i2bc.paris-saclay.fr (Y.Z.); gaelle.lelandais@i2bc.paris-saclay.fr (G.L.); genevieve.coste@i2bc.paris-saclay.fr (G.C.); claire.bouthier@i2bc.paris-saclay.fr (C.B.d.l.T.); esma.bentchikou@i2bc.paris-saclay.fr (E.B.); pascale.servant@i2bc.paris-saclay.fr (P.S.)
* Correspondence: fabrice.confalonieri@i2bc.paris-saclay.fr
† These authors contributed equally to this work.

Abstract: Numerous genes are overexpressed in the radioresistant bacterium *Deinococcus radiodurans* after exposure to radiation or prolonged desiccation. It was shown that the DdrO and IrrE proteins play a major role in regulating the expression of approximately twenty genes. The transcriptional repressor DdrO blocks the expression of these genes under normal growth conditions. After exposure to genotoxic agents, the IrrE metalloprotease cleaves DdrO and relieves gene repression. At present, many questions remain, such as the number of genes regulated by DdrO. Here, we present the first ChIP-seq analysis performed at the genome level in *Deinococcus* species coupled with RNA-seq, which was achieved in the presence or not of DdrO. We also resequenced our laboratory stock strain of *D. radiodurans* R1 ATCC 13939 to obtain an accurate reference for read alignments and gene expression quantifications. We highlighted genes that are directly under the control of this transcriptional repressor and showed that the DdrO regulon in *D. radiodurans* includes numerous other genes than those previously described, including DNA and RNA metabolism proteins. These results thus pave the way to better understand the radioresistance pathways encoded by this bacterium and to compare the stress-induced responses mediated by this pair of proteins in diverse bacteria.

Keywords: radioresistance/desiccation; transcriptional regulator; *Deinococcus radiodurans*; ChIP-seq; RNA-seq; bioinformatic analyses

1. Introduction

Deinococcus radiodurans is one of the most resistant bacteria to genotoxic agent exposure and desiccation isolated to date [1–4]. Unlike radiosensitive organisms, once exposed to huge γ-ray doses, or after prolonged desiccation, *D. radiodurans* is able to reconstruct an intact genome in a few hours from several hundred DNA fragments [5]. Many factors contribute to the radioresistance of *D. radiodurans*, including efficient DNA repair mechanisms [5–8], a condensed nucleoid limiting the dispersion of genome fragments after irradiation [9,10], and the protection of proteins against oxidative damage [11]. Thus, the exceptional ability of this bacterium to overcome severe DNA damaging conditions is described as a combination of active and passive mechanisms acting in synergy within the cell, enabling survival following these stresses.

The exposure of *D. radiodurans* to γ-rays, or its recovery from desiccation, results in a rapid upregulation of the expression of numerous genes [12,13], even if constitutively expressed genes are also involved in the mechanisms of radioresistance. In many bacterial species, expression of DNA repair genes is under the control of LexA, the repressor of the well-known SOS response (for review [14]). *D. radiodurans* encodes two LexA homologs

(LexA1 and LexA2) that undergo, as in *E. coli*, a RecA-dependent cleavage after DNA damage, but neither LexA1 nor LexA2 are involved in the induced expression of RecA [15,16]. In *Deinococcus*, the main response pathway to genotoxic conditions is regulated by the constitutively expressed metalloprotease IrrE [17,18] and the transcriptional repressor DdrO [19,20]. In vivo, the loss of function of IrrE completely abolishes the induction of the expression of numerous genes after exposure to ionizing radiation, and resulted in significant sensitivity of the strain to genotoxic conditions [19–25]. When cells are exposed to genotoxic stress conditions, IrrE cleaves the C-terminal region of DdrO [17,25,26], abolishing its DNA binding properties and leading to the expression of the genes repressed by DdrO [19]. Recent data suggest two different intracellular signals to induce the RDR regulon: (i) direct DNA damage [27] and (ii) a redox signaling pathway including zinc as a second messenger [25].

The DdrO protein is composed of two domains: an N-terminal helix-turn-helix (HTH) XRE DNA-binding domain, itself associated with a specific structural domain at the C-terminus required for protein dimerization and for DNA binding [28,29]. In vitro, IrrE-mediated cleavage removes the C-terminal 23 amino acid residues from DdrO [17,28].

The *ddrO* gene is essential for cell viability of *D. radiodurans* and *Deinococcus deserti* [17,20]. Interestingly, its prolonged depletion by a conditional deletion system induces, in *D. radiodurans*, an apoptotic-like response (DNA degradation, defects in chromosome segregation, and membrane blebbing) [20]. These results suggest that management of DNA damage can lead to cell survival or cell death. In *D. radiodurans*, these two responses are mediated by common regulators, IrrE and DdrO [20].

The IrrE/DdrO protein pair is highly conserved in *Deinococcus* species, and genes encoding IrrE/DdrO-like proteins are also present in other bacteria [30,31]. However, questions remain about the number of the genes that are directly or indirectly regulated by these two proteins. A 17-base-pair palindromic motif, designated as the Radiation/Desiccation Response Motif (RDRM), was identified in the promoter regions of several radiation-induced genes in different *Deinococcus* species, suggesting the existence of a conserved radiation/desiccation response (RDR) regulon [19,32,33]. The predicted RDR regulon of seven *Deinococcus* species consists of at least 14–24 genes, including numerous genes involved in DNA metabolism, such as *recA*, *ssb*, *gyrA*, *gyrB*, *uvrA*, and *uvrB*, but also *Deinococcus*-specific genes, such as *ddrA*, *ddrB*, *ddrC*, *ddrD*, and *pprA* [19]. Based on the presence of the RDRM located in the promoter region of the most highly upregulated genes by ionizing radiation and desiccation, 25 genes were predicted to belong to the RDR regulon in *D. radiodurans*: 24 genes by Makarova et al. 2007 [33] and *ddrC* [21]. It has been shown that, in vitro, *D. radiodurans* DdrO was able to bind to 21 predicted RDRM motifs [26] and in vivo, mutations within the RDRM sequence, in addition to the transient depletion of DdrO, induced the expression of several RDR regulon genes [20,21,34].

In this study, we mapped the DdrO regulon in *D. radiodurans* using two genome-scale approaches, i.e., ChIP-seq and RNA-seq analyses. These approaches were performed to identify additional DdrO target sites that were not predicted by previous in silico analyses. To our knowledge, we present here the first ChIP-seq analysis performed at the genome level in *D. radiodurans*. Alignments of DNA sequences extracted from ChIP-seq analysis were also performed to compare the consensus motif found in the predicted RDRM. As a prerequisite to a robust feature-mapping study, we resequenced our laboratory stock strain of *D. radiodurans* R1 ATCC 13939 to obtain an accurate reference for read alignments and gene expression quantifications.

Our results show 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, but also new genes associated with different metabolic pathways, involved in the translation or encoding of proteins of unknown function.

2. Material and Methods

2.1. Bacterial Strains; Plasmids; Oligonucleotides; Media

Bacterial strains and plasmids are listed in Table S1. The *E. coli* strain DH5α was used as the general cloning host, and strain SCS110 was used to propagate plasmids prior to transformation of *D. radiodurans* [35]. All *D. radiodurans* strains were derivatives of the wild-type strain R1 ATCC 13939. Transformation of *D. radiodurans* with PCR products, genomic or plasmid DNA was performed as previously described [6]. Strains expressing V5-tagged proteins were constructed by the tripartite ligation method as previously described [36]. The genetic structure and purity of the mutants were verified by PCR and sequencing. The sequences of oligonucleotides used for strain and plasmid construction are listed in Table S2. Chromosomal DNA of *D. radiodurans* was extracted using the NucleoSpin DNA Microbial Mini kit (Macherey-Nagel, Duren, Germany). *D. radiodurans* genomic DNA used to sequence the genome with Nanopore technologies was prepared by a lysis procedure involving a pretreatment of the cells with saturated-butanol in EDTA [37]. PCR amplification of DNA fragments, using plasmid or genomic DNA as a template, was performed using Phusion DNA polymerase (Thermo Scientific, Waltham, MA, USA).

D. radiodurans strains and derivatives were grown at 30 or 37 °C in TGY2X (1% tryptone, 0.2% dextrose, 0.6% yeast extract), or plated on TGY1X containing 1.5% agar, and *E. coli* strains were grown at 37 °C in Lysogeny Broth. When necessary, media were supplemented with the appropriate antibiotics used at the following final concentrations: kanamycin, 6 µg/mL; chloramphenicol, 3.5 µg/mL; hygromycin 100 µg/mL; spectinomycin, 90 µg/mL for *D. radiodurans*; chloramphenicol, 25 µg/mL; spectinomycin 50 µg/mL for *E. coli*.

2.2. D. Radiodurans R1 Sequencing; Assembly and Annotation

Purified *D. radiodurans* genomic DNA from strain R1 ATCC 13939 (laboratory stock) was sequenced using Illumina NextSeq v. NS500446 (High-throughput sequencing facility of I2BC, Gif sur Yvette, France), yielding 10.7×10^6 75 nt paired-end reads. This dataset was subsequently assembled with SPAdes (v3.13.1) St Petersburg University, Russia [38], with the–'careful' option set to reduce the number of mismatches and short indels, and produced a total of 3,255,298 nt. in 136 contigs (N50: 149391).

In parallel, *D. radiodurans* genomic DNA was also sequenced with Oxford Nanopore Technologies GridION (v. GXB02022–19.12.6) (High-throughput sequencing facility of the I2BC, Gif sur Yvette, France), yielding a total 10.87×10^9 nucleotites in 946,434 long reads (median size: 6857 nt). This reads pool was further filtered with filtlong (v. 0.2.0) (https://github.com/rrwick/Filtlong) (accessed on 10 September 2021), retaining only reads longer than 2000 nt, aligned to reference Illumina reads (see above) and totaling $\sim 5 \times 10^9$ nt. This reads dataset was used as an input to Canu (v1.8) [39] for reads correction and assembly with default parameters, producing 4 low quality contigs, totaling 3,578,820 nt.

These latter sequences were improved by mapping SPAdes-produced high quality contigs on them with BWA mem (v. 0.7.9a-r786) [38]. SPAdes contigs alignments on nanopore reference sequences were extracted with samtools mpileup (v. 1.8) [40] and variants were called with bcftools call (v. 1.10.2) [41,42] using prior probability set to 1 as settings. Amino acid sequences of predicted genes were searched for similarity with BLASTP [43] to sequences from two other available complete sequences of *D. radiodurans* R1 strain (GCA_000008565.1 and GCA_001638825.1). Structural RNAs were mapped on the genomic sequences with the same reference genomes using the BLASTN tool.

CDS prediction was performed on the final assembled sequences using Prodigal (v. 2.6.3) using single mode, translation Table 11 [44] as settings. Amino acid sequences of predicted genes were searched for similarity with BLASTP [43] to sequences from two other available complete sequences of *D. radiodurans* R1 strain (GCA_000008565.1 and GCA_001638825.1). Structural RNAs were mapped on the genomic sequences with the same reference genomes using the BLASTN tool.

2.3. Time Course Experiment

2.3.1. DdrO Depletion

The p17235 and p17236 shuttle vectors (*E. coli*/*D. radiodurans*) code for a spectino-mycin resistance gene, in addition to a *repU* gene encoding the RepU protein essential for plasmid replication in *D. radiodurans*. The p17235 and p17236 plasmids contain the wild-type *repU* gene or a *repU*$_{Ts}$ gene encoding a thermosensitive protein (RepU$_{Ts}$), respectively. To construct the p17238 plasmid used for conditional expression of *ddrO*, the native genomic *ddrO* gene was amplified by PCR using the NE28-NE29 primers and the PCR product was cloned into the p17236 plasmid between the *Bam*HI/*Not*I sites. The same strategy was used to construct the p17237 plasmid. Following transformation of *D. radiodurans* strain GY14125 (non-homogenotized Δ*ddrO*Ω*cat*) with p17236 or p17237 plasmids, both expressing *ddrO* gene, the transformants were streaked several times on plates supplemented with chloramphenicol and spectinomycin, and the complete deletion from all chromosome copies of native *D. radiodurans ddrO* was analyzed by diagnostic PCR.

Δ*ddrO* strains complemented by *ddrO* expressed, under its own promoter region, from a plasmid with wild-type or thermosensitive (prepU$_{ts}$) replication, were grown at a permissive temperature (30 °C) with spectinomycin and chloramphenicol. Cells were diluted in fresh medium with antibiotics and grown at permissive temperature (30 °C) and cells in exponential growth (A$_{650nm}$~0.5) were harvested by centrifugation, washed two times with TGY2X and reused to A$_{650nm}$ = 0.1 in fresh medium without antibiotics. Then, the temperature was shifted to 37 °C (non-permissive temperature for the thermosensitive replication plasmid). At 1, 4, 6, 8, 16, and 24 h, aliquots of 20 mL were removed for fluorescence microscopy and transcriptome analysis or Western blot analysis.

2.3.2. RNA Extraction, cDNA Library Construction, and Sequencing

For each aliquot, total RNA was isolated using the Fast RNA Pro Blue Kit (MP Biomedicals, Illkirch-Graffenstaden, France) and the FastPrep-24 instrument, according to the manufacturer's protocols. Extracted RNA was rigorously treated with TURBO DNA-free (Invitrogen, Waltham, MA, USA), according to the manufacturer's instructions and the absence of DNA genomic contamination was checked by quantitative PCR (qPCR). The quality and quantity of treated RNA were analyzed using a DeNovix DS-11 spectrophotometer (DeNovix Inc., Wilmington, DE, USA) and the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA) with an RNA integrity number ≥6 for cDNA library preparation. The rRNA depletion and Illumina libraries were made following the Illumina protocol (High-throughput sequencing facility of the I2BC, Gif sur Yvette, France). The *cDNA* samples were sequenced using Illumina NextSeq v. NS500446 (High-throughput sequencing facility of the I2BC, Gif sur Yvette, France), yielding, on average, 22.8 × 10^6 50 nt. paired-end reads (±6.8 × 10^6 reads).

2.3.3. RNA-Seq Data Analysis

Read sequences were mapped on our reference genome sequence with BWA mem (v. 0.7.9a-r786) using default settings, and coverage values of all genomic features were computed with the bedtools "coverage" command (v2.17.0) [45]. RNA differential gene expression analysis was performed with the DESeq R-package v. 1.39.0 [46].

2.4. Western Blot Analysis

The protein extractions and Western blot analyses were performed as previously described [20]. The membranes were incubated overnight at 4 °C with a 1:5000 dilution of monoclonal rabbit anti-HA antibodies (Sigma-Aldrich, Saint-Louis, MO, USA) or 1:5000 dilution of monoclonal rabbit anti-FLAG antibodies (Sigma-Aldrich, Saint-Louis, MO, USA).

2.5. Chromatin Immunoprecipitation (ChIP)

Exponentially growing *D. radiodurans* cells (100 mL, A_{650nm} = 0.7) expressing DdrO-V5 or native DdrO protein were crosslinked with 1% formaldehyde in TGY2X medium for 25 min at 30 °C with continuous shaking. Crosslinking reactions were quenched by the addition of 125 mM glycine for 15 min. Cells were harvested by centrifugation (4000× *g*, 10 min, 4 °C), washed twice with cold Tris Buffer Solution (TBS, 50 mM Tris, 100 mM NaCl pH 7.5), and then resuspended in 3 mL of lysis buffer (160 nM NaCl, 20 mM, Tris-HCl pH7.5, 1 mM EDTA, protease inhibitor cocktail (Roche)). Cells were disrupted and DNA sheared using a One Shot Cell Disruptor (CellD SARL) to an average size of 100–300 bp (2 rounds of 2.4 kbar). Insoluble material was removed by centrifugation at 20,000× *g* for 10 min at 4 °C and the supernatant was collected in a sterile microcentrifuge tube. Then, 500 μL of supernatant fluid was added to 25 μL of pre-incubated protein G magnetic beads (ChIP-Adembeads ChIP-Adem-Kit, Ademtech SA, Pessac, France) with 5 μg of anti-V5 rabbit polyclonal antibody (ab9116, Abcam, Cambridge, UK) in IP buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton 100, protease inhibitor cocktail (Roche). After overnight incubation at 4 °C with rotation, the immuno-precipitates were washed 5 times with washing buffers (ChIP-Adembeads ChIP-Adem-Kit, Ademtech SA, Pessac, France).

Immune complexes were eluted in 200 μL of elution buffer. The eluted samples (20 μL) were saved for control Western blots, and the remainder was incubated for 2 h at 37 °C with shaking with 100 μg/mL Proteinase K. Then, the supernatant was incubated overnight at 65 °C to reverse crosslinking with 100 μg/mL RNAse A. The DNA was purified using the PCR Clean-up kit (Macherey-Nagel, Duren, Germany). Three independent ChIP experiments were performed for "IP" samples.

2.6. ChIP-Seq

Raw FASTQ files were obtained from sequencing three "IP" samples comprising 19, 7, and 3 × 10^6 sequences, respectively, in addition to the "Mock" (19 × 10^6 sequences) and the "Input" (13 × 10^6 sequences) samples. The quality score was verified with FASTQC software (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) (accessed on 10 September 2021) and Illumina adaptor sequences were removed with Cutadapt software (https://cutadapt.readthedocs.io/en/stable) (accessed on 10 September 2021). Sequence alignments on the genomic sequence were performed with Bowtie2 software (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) (accessed on 10 September 2021). Output SAM files were converted and indexed into BAM files, using the Samtools software (http://www.htslib.org) (accessed on 10 September 2021). They were used both for visualization with IGV, and additional conversion into BED files with Bedtools software (https://bedtools.readthedocs.io/en/latest/index.html) (accessed on 10 September 2021) providing the input file format required by bPeaks programs (Available online: https://cran.r-project.org/web/packages/bPeaks/index.html) (accessed on 10 September 2021) to perform peak calling. Searches for conserved motifs were performed by MEME and FIMO (https://meme-suite.org/meme) (accessed on 10 September 2021) with a Match *p*-value < 1.0×10^{-4}. Prediction of *E. coli*-like gene promoter elements and transcription start sites in gene promoters was carried out using BPROM (http://www.fruitfly.org/seq_tools/promoter.html) (accessed on 10 September 2021) [47]. To sort data from Chip-seq and RNA seq and to integrate them with the conserved motifs found by MEME and FIMO, we used an in-house script (File S1) that defines, without a priori knowledge, different lists of candidate genes to be DdrO targets, and hence provides detailed information about the process, which was applied to obtain the results presented in the main text and Supplementary Materials files.

2.7. Western Blot Analysis of RDR Tagged-Proteins

Exponentially growing bacteria (15 mL, A_{650nm} = 0.3), grown at 30 °C, were exposed to 1 or 5 μg/mL mitomycin C. After 3 h at 30 °C with continuous shaking, cells were

harvested by centrifugation at 4 °C and the pellets washed with 1X cold saline-sodium citrate (SSC) buffer. Then, the bacteria pellets were re-suspended in 150 μL of SSC 1X with 0.4 mM protease inhibitor cocktail (Roche) and cells disrupted with a FastPrep Instrument using 0.1 g of glass beads (500 μm) and four pulses of 30 s. Cell debris were removed by centrifugation at 20,000× *g* for 10 min at 4 °C, and the supernatant fluid collected and placed in sterile microcentrifuge tubes. Protein concentrations were determined by Bradford assay (Bio-Rad). Proteins were subjected to electrophoresis through a 12% Glycine SDS-PAGE gel (Mini-PROTEAN TGX Stain-Free Precast gel, Bio-Rad, Hercules, CA, USA) and transferred onto a polyvinylidene difluoride (PVDF) membrane (GE Healthcare, Chicago, IL, USA). All these experiments were performed three times as previously described [20] with a 15,000 dilution of anti-V5 rabbit primary antibody (Abcam, Cambridge, UK) or with a 1:5000 dilution of monoclonal rabbit anti-HA antibodies (Sigma-Aldrich, Budapest, Hungary).

2.8. Sensitivity Assay to DNA-Damaging Agents Mitomycin C and UVC

Bacteria were grown in TGY2X liquid medium at 30 °C to an $A_{650nm} = 1$ and sequential dilutions of cells were spotted on TGY plates supplemented (or not) with mitomycin C (60 ng/mL and 80 ng/mL at final concentration), exposed (or not) to UVC at a dose rate of 3.5 J/m^2/s.

2.9. Deposition of Sequences and of Expression Data

The complete sequence and annotation of the genome were deposited with GenBank under accession numbers CP068791, CP068792, CP068793, and CP068794. The complete high-throughput sequence data were deposited with the Gene Expression Omnibus (GEO) data bank under accession number GSE175875 (RNA-seq and ChIP-seq).

3. Results

3.1. Genome Sequencing

Two genome sequences of the *D. radiodurans* strain R1 are available in databases [48,49]. The genome is distributed over four replicons: two chromosomes, one megaplasmid, and a plasmid (Table 1). A nucleotide polymorphism between the two complete genome sequences was reported, in addition to several insertions, deletions, or substitutions frequently found in bacterial genomes [49]. To promote accurate RNA-seq and ChIP-seq analysis, and for searching for conserved binding motifs for the DdrO protein, we sequenced the *D. radiodurans* genome of strain R1 ATCC 13939 maintained in our laboratory. We opted for Illumina NextSeq Oxford sequencing coupled with Nanopore Technologies GridION to unambiguously locate the repeated elements that may misassemble short sequences in size. Merging both sets of sequences produced an ensemble of four high quality contigs, totaling 3,578,820 nt, with a 450-fold average coverage. Among the 3230 predicted genes, 3147 encode proteins.

As shown in Table 1, the size of chromosome 2 and the two plasmids, in addition to the number of CDS encoded by *D. radiodurans* deduced from our sequence, are closer to those published by White et al. [48] than these published by Hua [49]. The large sequence insertions revealed in the more recent release were not found here. However, because the sequence published by White et al. [48] contains many errors, the degree of identity of genes was better with the genome sequence published by Hua [49], with a higher percentage of genes found between these two releases when a threshold of 90% of maximum bit score was applied (Table 1). The sequence origin for each chromosomal element and plasmid was adjusted to correspond to the genome sequence of White et al. [48]. In addition to the orthologs of CDSs with both previously sequenced genomes as listed in Table S3, numerous genes or pseudo genes from different transposon families were found.

Table 1. Size of each replicon, %GC, and CDS content in *Deinococcus radiodurans* strain R1 ATCC 13939 genome sequence. Percentage of pairs of orthologs found in the genome sequence of our strain when compared to the two other strains' R1 genome sequence when a threshold of 90% of maximum bit score was applied.

	Replicon	ID	Percentage of Pairs of Orthologs *	CDS	Total CDS	Nt.	GC%
Hua and Hua (2016)	Chr 1	CP015081	96.43	2523		2,646,742	67.07
	Chr 2	CP015082	96.02	352	3079	433,133	66.77
	Megaplasmid	CP015083	81.05	153		203,183	62.98
	Plasmid	CP015084	62.75	51		61,707	56.55
White et al., (1999)	Chr 1	DRA1	84.37	2629		2,648,638	67.01
	Chr 2	DRA2	85.05	368	3181	412,348	66.69
	Megaplasmid	DRA3	77.24	145		177,466	63.19
	Plasmid	DRA4	61.54	39		45,704	56.15
This work	Chr 1	DRO	100	2594		2,644,251	67.08
	Chr 2	DRO_A	100	364	3147	412,138	66.65
	Megaplasmid	DRO_B	100	148		177,322	63.21
	Plasmid	DRO_C	100	41		45,508	56.26

* Percentage of pairs of orthologs (at a threshold of 90% max bit score) found in the genome sequence of our strain when compared to each replicon from the two other strain R1 ATC 13989 genome sequence releases.

3.2. In Vivo Identification of DdrO Binding Sites by ChIP-Seq Assays

To localize in vivo the chromosomal regions bound by the DdrO protein, we constructed the GY 18218 strain expressing a V5-tagged DdrO protein, in all the genome copies, from the native promoter region of *ddrO* (Figure S1). Cells expressing the recombinant protein, tagged at its C-terminal end, displayed the same growth rate as the wild-type strain and the expression of DdrO-V5 did not affect the resistance of the strain to DNA damaging agents (mitomycin C and UV) (Figure S1). These results demonstrate that DdrO-V5 protein is functional and remained cleavable by IrrE under stress conditions.

D. radiodurans GY 18218 and R1 strains were grown to mid-log phase and ChIP-seq was performed on DNA precipitated by the ChIP grade anti-V5 antibody. The Input sample (chromosomal DNA of the GY 18218 strain), the Mock sample (immuno-precipitated (IP) DNA of the wild-type strain), and three replicates of the DdrO-V5 IP sample were used to prepare sequencing libraries. The DNA regions over-represented in the DdrO-V5 IP sample and corresponding to potential binding sites for DdrO-V5 were identified using the bPeaks program [50].

A total of 136 ChIP-enriched peaks were found, mainly (110/136) within intergenic regions of genes encoding proteins (Figure 1A and Table S4), whereas 26 peaks were intragenic or found at the vicinity of genes coding for tRNA. Significant peaks, as illustrated for five genes, *ddrA*, *ddrB*, *ddrC*, *gyrA*, and *gyrB* (Figure 2), were identified in the promoter region of 18 of 25 genes reported as belonging to the RDR regulon (Figure 2 and Figure S2) [21,26,33].
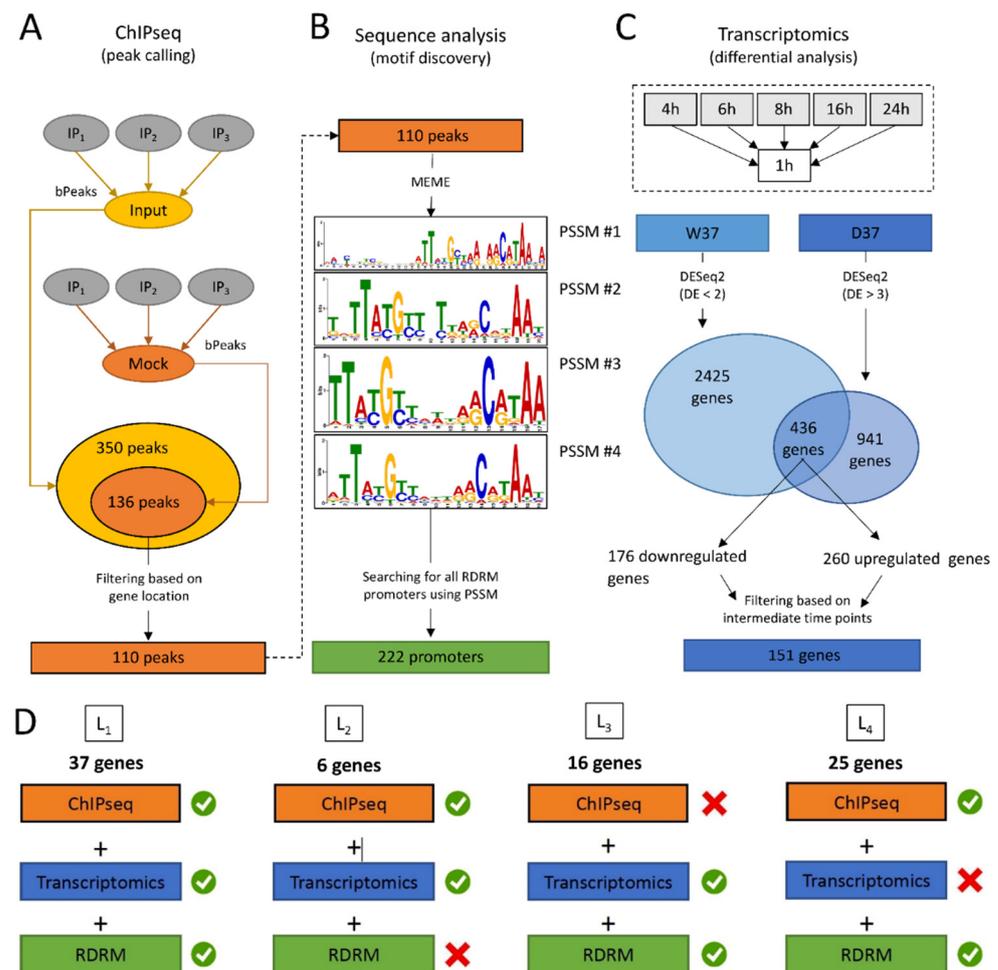
Figure 1. Overview of the computational strategy used to integrate omics data and identified candidate genes for inclusion in the DdrO regulon. (**A**). ChIP-seq analyses, i.e., defining the genomic regions for which interactions between DNA and the DdrO protein were observed. The three IP replicates were compared to the Input and Mock controls. Peaks identified in both comparisons were retained. An additional filter was applied to focus on only the peaks located in intergenic regions. (**B**). Sequences of the peaks identified in (**A**) were used to search for over-represented DNA motifs, applying the MEME program. Four position specific scoring matrixes (PSSMs) were retained, because of their redundancies. PSSM were used as inputs for the FIMO program, scanning sequences between −500 and +100 of all annotated CDS. Positive matches were retained and are referred to as "RDRM promoters". (**C**). RNA-seq data was used to identify differentially expressed genes, comparing each time point (4, 6, 8, 16, and 24 h) to the first (1 h). In the W37 strain, genes identified as differentially expressed in less than two comparisons were selected (DE < 2), whereas in the D37 strain, genes identified as differentially expressed in more than three comparisons were selected (DE > 3). Common genes from the two lists were retained and an additional filter was applied to focus further analyses on only these genes which exhibited differential expression (up- or downregulation) at intermediate times, i.e., 6, 8, and 16 h. (**D**). Results obtained in (**A–C**) were integrated to define four lists in the DdrO regulon. The first list is comprised of genes for which (i) a peak was detected upstream of the gene location (ChIP-seq results), (ii) specific differential expression was observed in the D37 strain (transcriptome results), and (iii) an RDRM motif was found in the gene's promoter region (sequence analysis). Other lists (L2-L3-L4) matched with only two criteria (green tick mark). These genes are worth considering as potential targets for DdrO.
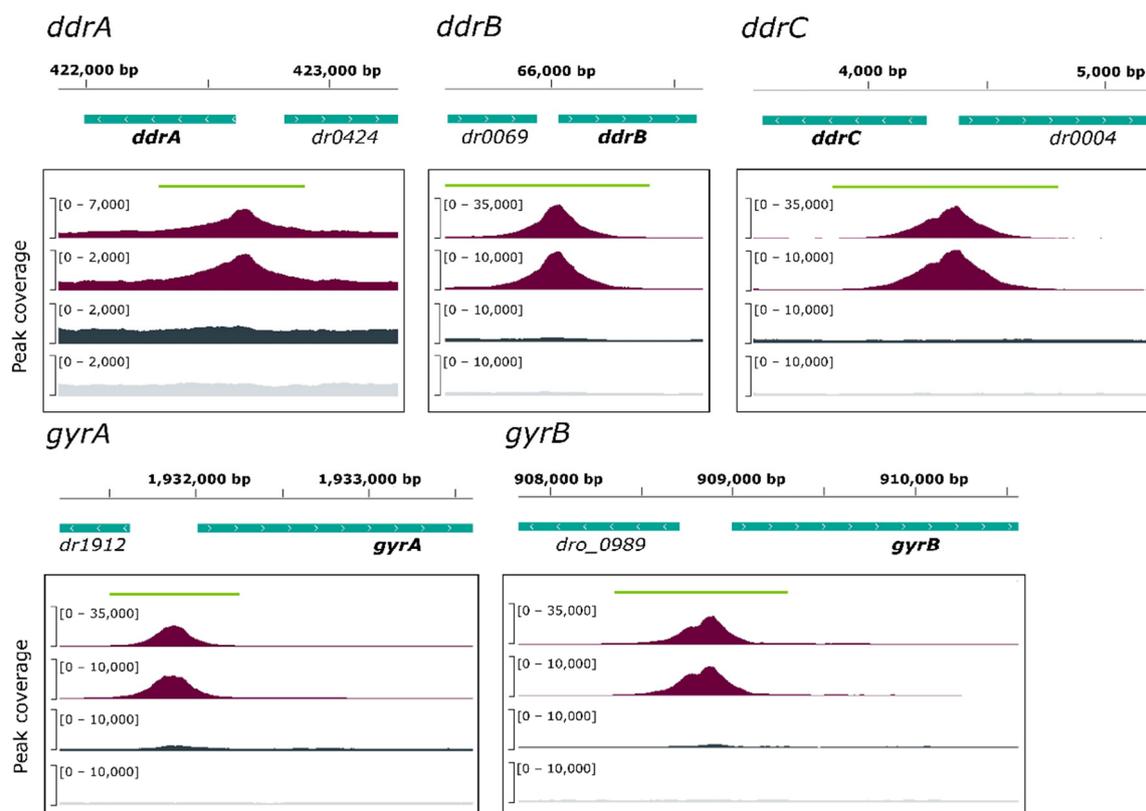
Figure 2. Visualization, through IGV, of the binding peaks obtained from genome analysis. Tag density profiles are illustrated for 2 IP (purple), the Input (dark grey), and the Mock (light grey) for five known DdrO-regulated genes: *ddrA*, *ddrB*, *ddrC*, *gyrA*, and *gyrB*. The green lines indicate the size of each peak identified by bPeaks. Genes are represented by green boxes, their location on either strand is indicated by > (strand +) and < (strand −). The genomic coordinates of each locus are indicated on the *X*-axis. *dro_0898* is split into two CDS (DR0903-DR0904) in the White et al. annotation [48].

A careful inspection of DdrO-V5 IP tag density through the IGV program of the seven missing genes showed that a lower coverage of reads was observed at the promoter region of *mutL*, and small peaks were observed in the promoter regions of *recQ* and *sbcD* that fell below the threshold used for peak detection with the bPeaks program. No peaks were detected in the region upstream of the *hutU*, *irrI*, *frnE*, and *rsr* genes. To identify candidate binding sites of the DdrO protein, the nucleotide sequences of the ChIP peaks, between 151 and 1401 in length, were compared using MEME [51], to search for palindromic or non-palindromic motifs with an occurrence of one motif per sequence or any number of repetitions (Figure 1B). A total of 41 peak sequences, located in the promoter regions, contained a conserved DNA motif close to the RDRM sequence, with some loci containing two motifs (Table S4), as illustrated for DdrA (Figure 2). A broad peak was observed in its promoter region because it contains two RDRM (Table S4). Interestingly, based on ChIP-seq data, the number of RDRM reported here is larger than that predicted by previous in silico analyses [19,33]. No other conserved sequence pattern was found, with the exception of the predicted core promoters, either from the 41 ChIP peaks sequences or from the other sequences lacking an RDRM. However, a degenerate RDRM may not have been detected due to the threshold used for these bioinformatic analyses. Altogether, these results confirmed in vivo the role of the RDRM for DdrO binding to the *D. radiodurans* genome. Independently, we investigated whether an RDRM was found in the promoter regions of other genes encoded by *D. radiodurans*. We monitored, with FIMO, their presence in a set of sequences covering the regions located between −500 and +100 nucleotides from the start of translation of all the 3147 CDS encoded by *D. radiodurans*. A total of 222 putative RDRM-like sequences (Figure 1B, Table S5) were found, including the 41 detected

by MEME and eight other potential sites, that were not detected by MEME, but six of which were located far from the start of coding sequences and outside the ChIP-seq peaks.

Based on ChIP-seq results, 89 genes, sometimes included in operons, may be regulated by DdrO, but many enriched peaks were located within the intergenic region of divergently transcribed genes. It is possible that only one of the two divergent genes may be under the control of DdrO.

3.3. Transcriptome Analysis of D. radiodurans in Response to the Depletion of DdrO

In parallel, to further characterize the RDR regulon in *D. radiodurans*, we compared transcriptome profiles of cells expressing, or not, DdrO. Because DdrO is essential for cell viability [17,20], we used a conditional gene inactivation system [20,52]. In this system, $\Delta ddrO$ cells expressed the DdrO protein under control of its own promoter region at 30 °C from a temperature-sensitive ($repU_{Ts}$) replication vector (Figure 3A) [20]. Shifting the culture to 37 °C, a non-permissive temperature, resulted in an inability of the plasmid to replicate during successive cell divisions, leading to the depletion of DdrO, in contrast with a derivative of this expression vector, containing the wild-type $repU^+$ gene, that did not cause depletion of DdrO at 37 °C [20,52]. The $\Delta ddrO/ddrO^+$ ($prepU_{Ts}$) and $\Delta ddrO/ddrO^+$ ($prepU+$) strains grown at 37 °C are denoted D37 and W37, respectively. Under our experimental conditions, the number of cells carrying the $repU^+$ vector is proportional to the increase in cell mass at 37 °C without a selective antibiotic (Figure 3B). In contrast, the number of cells carrying the $repU_{TS}$ vector stopped increasing and remained stable over 24 h (Figure 3B). The growth curves of both strains exhibited a comparable doubling time over 6 h. However, after this time lapse, the growth of the D37 strain also stopped, coinciding with the stress triggered by the depletion of DdrO (Figure 3B).

In a first attempt, we analyzed the effect of DdrO depletion on the expression of DdrD, DdrO, PprA, and RecA proteins belonging to the RDR regulon. For this purpose, we used derivatives of the strain $\Delta ddrO$ ($prepU_{Ts}:ddrO^+$) expressing DdrO-FLAG, PprA-HA, DdrD-HA, or RecA-HA tagged proteins. Depletion of the DdrO repressor, in cells grown at the non-permissive temperature, resulted in the complete loss of DdrO after 4 h at 37 °C (Figure 3C) and an increasing amount of PprA, DdrD, and RecA proteins during the kinetics (Figure S3).

In a second step, the kinetics of gene expression changes induced by DdrO depletion were analyzed for both strains, from three independent cultures and at six time points (1, 4, 6, 8, 16, and 24 h) (Figure 4A). RNA sequencing data was performed from 36 samples, corresponding for each sample to an average sequencing depth of 647-fold the genome sequence.

A two-fold change in expression threshold for the ratio in these experiments was applied, together with a *p*-value < 0.01. Principal component analysis (PCA) confirmed that the transcriptome of the three biological replicates was clustered at each time point, showing the reproducibility of the experiments and the transcriptome patterns evolved as cells progressed through the time course of the experiment (Figure 4B). The datasets are separate, in addition to the PC1 and the PC2 levels, which together explained approximately 75% of the variance.

To compare the transcriptome, the 1 h time point was used as the reference, giving time for the genome to stabilize its expression after shifting the temperature. We first compared, for each strain, the deregulation of all genes along the time course. The results of the differential expression for all genes in W37 and D37 are presented in Figure 5 and Tables S4 and S5. After an incubation of 24 h at 37 °C, numerous genes were deregulated, as 2129 unique genes i.e., 67.7% of all genes in the W37 strain, and 2330 unique genes, i.e., 74% in the D37 strain, were up- or downregulated at a minimum of one time point during the time course, showing that a cascade of cell regulation occurred into each strain over 24 h. Moreover, genes reported in one time point were often found in the following time point. From the set of 350 and 358 regulated genes in W37 and D37, respectively, during a time point between 1 and 4 h, 159 (45%) in W37 and 195 (54%) in D37 remained

regulated during all time points (Table S6 and Figure 5). After 6 h at 37 °C, although most of the cells in the D37 strain lost the thermosensitive plasmid, 679 and 1011 genes were deregulated in W37 and D37, respectively, with 501 genes shared between them (Table S6), and several upregulated genes found in W37 were downregulated in D37. These results showed that, rapidly after the temperature shift, the expression patterns of W37 and D37 changed differentially and the depletion of DdrO in D37 directly or indirectly deregulated the expression of other genes.
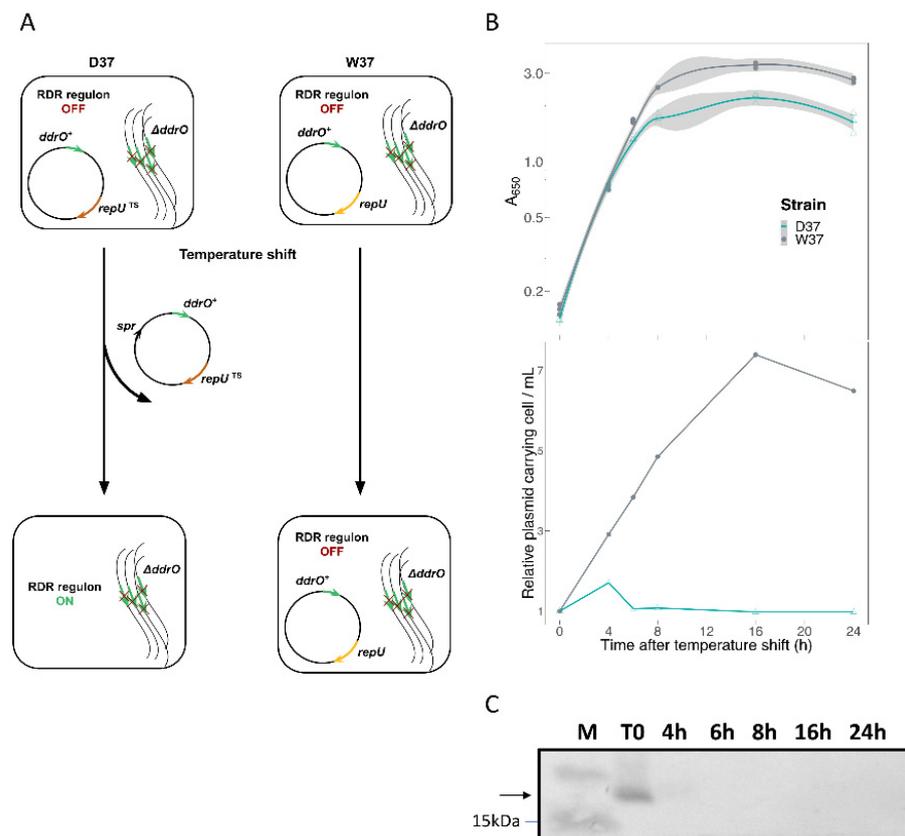

Figure 3. Loss of the *repU*$_{TS}$ vector in a chromosomal Δ*ddrO* context. (**A**). Experimental design. Expression of RDR regulon genes was induced at 37 °C when the thermosensitive replicative vector expressing DdrO could no longer replicate. (**B**). Growth parameters of W37 (grey line) and D37 (blue line) strains, and relative stability of the *repU*$_{Ts}$ or the *repU*$^+$ replication vectors expressing DdrO at 37 °C during a 24 h period. The A$_{650nm}$ values of the cultures were measured in 3 independent experiments (confidence interval for each smoothed curve is indicated in light grey area). To calculate vector stability, samples were removed at the indicated times for plating at 30 °C on media with or without spectinomycin. (**C**). Western blot analysis of recombinant DdrO-FLAG protein depletion in D37 at 37 °C. At each time point, aliquots of cells were removed and cell extracts (15 μg of proteins) were subjected to SDS-PAGE and analyzed by Western blot with anti-FLAG antibodies. The indicated times are relative to the initial temperature shift time point (0 h).
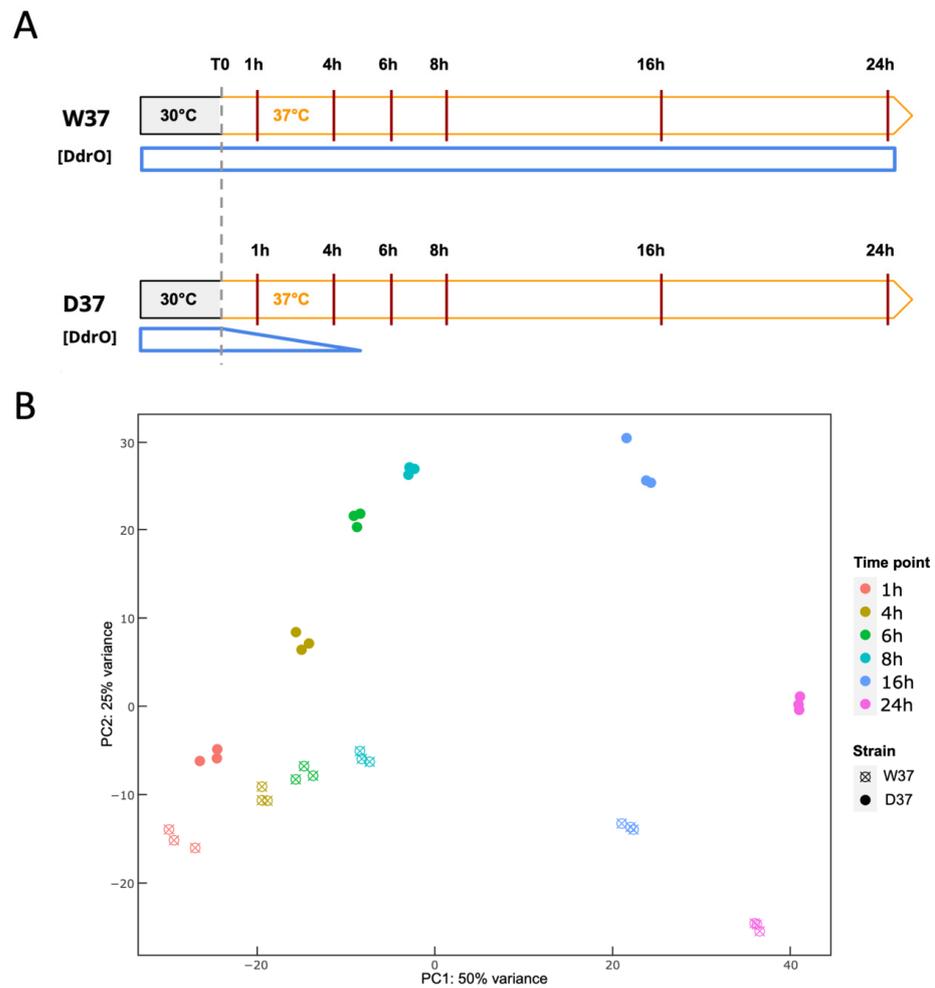
Figure 4. Transcriptomic analysis of DdrO depletion. (**A**). Schematic representation of the transcriptome time course. D37 and W37 cells were first cultivated at 30 °C and exponentially growing cells were rapidly transferred to 37 °C. The six time points examined are indicated by vertical lines. The amount of DdrO during the time course in both strains is indicated by blue plots. (**B**). Principal component analysis of all D37 and W37 samples after temperature shift. Each replicate was plotted as an individual data point. The indicated times are relative to the temperature shift time point (0 h).

To confirm the loss of the *prepU*$_{Ts}$ plasmid at 37 °C at the transcriptome level, we investigated the expression profiles of *ddrO* and *spr* encoding resistance to spectinomycin. As shown in Table S7, the *spr* gene was downregulated in D37, confirming the loss of the *prepU*$_{Ts}$ plasmid in growing cells at 37 °C (Figure 3), whereas the *ddrO* gene was upregulated in D37. When Δ*ddrO*/*ddrO*$^+$ strains were constructed, only the CDS encoding the DdrO protein was deleted from the genome. Therefore, the upstream region, containing the promoter and the 5′UTR region of *ddrO*, is duplicated, one located on the chromosomal locus, the second on the plasmid. The RDRM is located in the *ddrO* gene 153 nucleotides upstream of the ATG in the vicinity of the predicted promoter. The 5′UTR reads density profiles of *ddrO* were very low in the W37 strain, but were augmented in the D37 strain as soon as the cell lost the plasmid, supporting the findings of previous studies that show that DdrO regulates the expression of its own gene [17,26] (Figure S4).
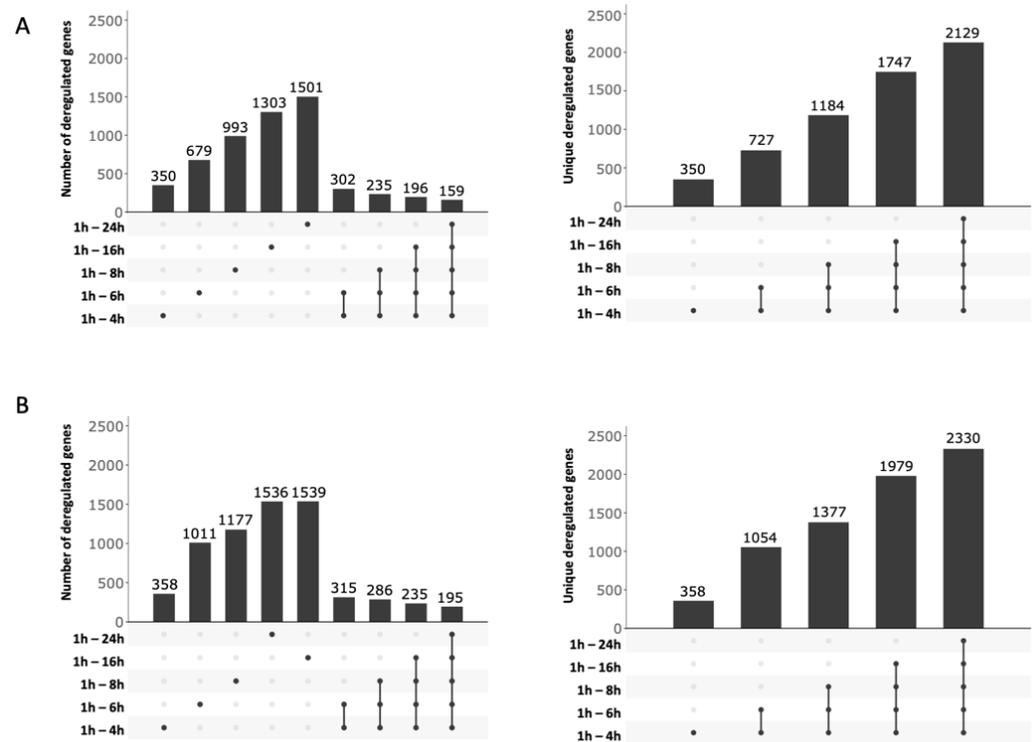
Figure 5. Evolution of deregulated genes from W37 (**A**) and D37 (**B**) strains after the temperature shift. The left panels show the number of differentially expressed genes at each time point (see Table S6) and the number of common genes among indicated time intervals (illustrated by solid black points, linked by black lines). The right panel shows the increasing number of unique genes that are differentially expressed at a minimum of one time point after the temperature shift. Number of deregulated genes ($|$FC$| \geq 2$, p-value ≤ 0.01).

Twenty genes of the RDR regulon were upregulated in the D37 strain, often from the beginning of the experimental temperature shift, with increasing fold changes as cells progressed through the time course. The *dr2256* gene encoding a transketolase, in addition to the *ddrF* and *ssb* genes, were also upregulated later (6 or 8 h, Table S7), and the *uvrA* gene (*dr1771*) was upregulated only after 16 h at 37 °C. The other genes, such as *uvrD* (*dr1775*) and *irrI* (*dr0171*), were not upregulated or only changed in a late stage of the experiment. In contrast, *drA0151*, encoding the first gene of the *hut* operon, was strongly downregulated in the W37 strain and in the D37 strain, but this gene was not reported as being under the control of the DdrO/IrrE proteins [24].

We also wondered if other *D. radiodurans* genes displayed a transcriptome pattern comparable to the RDR regulon genes. For this purpose, we selected genes as differentially expressed (DE) in the W37 strain in two or fewer comparisons (DE $\leq$ 2), and in more than three comparisons in the D37 strain (DE > 3), considering most of the profiles exhibited by the predicted RDR genes (see Methods, Table S4). A total of 436 genes displayed similar transcriptome profiles (Figure 1C, Table S8), reduced to 151 genes when an additional filter was applied to focus on genes which exhibited differential expression (up- or downregulation) at only intermediate time points, i.e., 6, 8, and 16 h. A total of 260, of the 436 identified genes, were upregulated in the D37 strain (Table S9), and 60% of these were distributed into four functional categories. A total of 31 genes encoding proteins involved in DNA metabolism, and including most of the previously predicted RDR regulon, were found, but also the *recG* and *recO* genes encoding a primosomal protein N′, in addition to two DNA polymerase III subunits (DR0001 and DR0507), *polA*, two putative helicases (DR0837 and DR1572), *ligA*, *mutS*, and *recN* (Figure 6). Interestingly, several genes encoding transcription factors involved in stress responses, such as LexA1, LexA2, or the Phage Shock Protein A

PspA, were also upregulated in D37 (Figure 6). Therefore, several regulatory networks were likely triggered in response to induction of the RDR regulon.
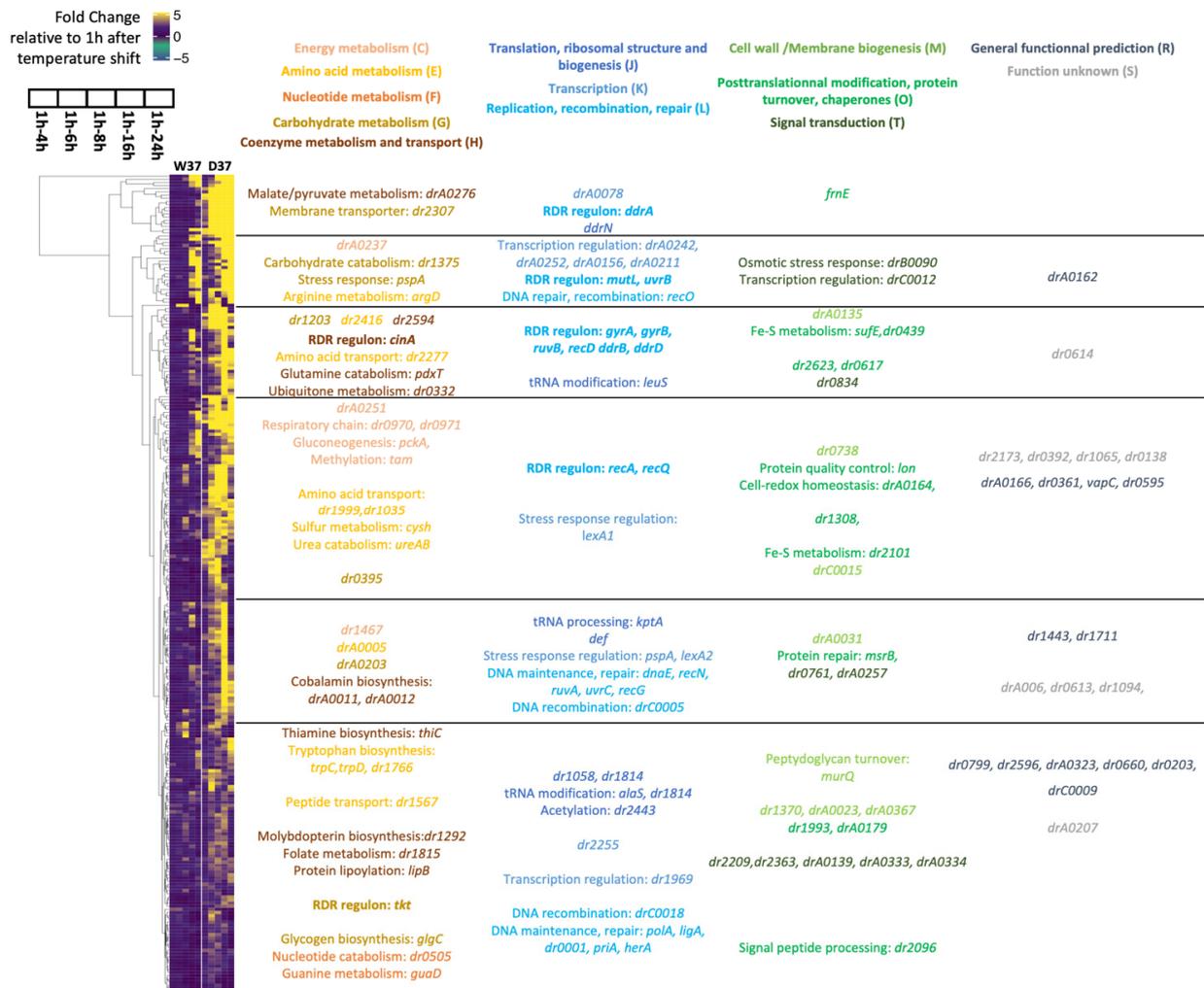

Figure 6. Hierarchical clustering of genes specifically upregulated in response to DdrO depletion. The 260 genes differentially expressed in the time periods 1–6 h and 1–16 h in the W37 strain in two or fewer comparisons (DE ≤ 2) and in more than three comparisons in the D37 strain (DE > 3) were hierarchically clustered according to their temporal expression. Only several genes representing some COG categories are shown.

In addition, 15 genes encoding putative proteases and peptidases or regulators of protease activity, 15 genes coding for ABC transporters, permeases, and efflux components, and >100 genes coding for uncharacterized proteins or of unknown function were also deregulated with similar expression patterns (Figure 6).

We also investigated the presence of downregulated genes during DdrO depletion. Using the same settings, 176 downregulated genes were found (Figure S5 and Table S10) to be widely distributed in the different COG categories, with a transcriptomic repression mainly beginning at 6 h when D37 cells are depleted in DdrO.

Our RNA seq analysis exhibited 436 deregulated genes, but the overall up- or down-expression of these genes may be a consequence of a cascade of regulation occurring when cells lost the *ddrO* gene.

3.4. Integration of the Data: The DdrO Map in D. radiodurans

To map the DdrO regulon, we integrated the results obtained in ChIP-Seq, RNA-seq, and motif research. When candidates from these three experiments were compared,

37 genes met all three criteria, or 42 considering that several genes are included in putative operons (Table 2), such as the *cinA-ligT-recA* operon previously described by Makarova et al. [33], and 47 other genes met two criteria (Figures 1D and 7). A consensus motif was searched for by MEME from all of the RDRM sequences listed in Table 2. This motif is consistent with those previously described [20,34].
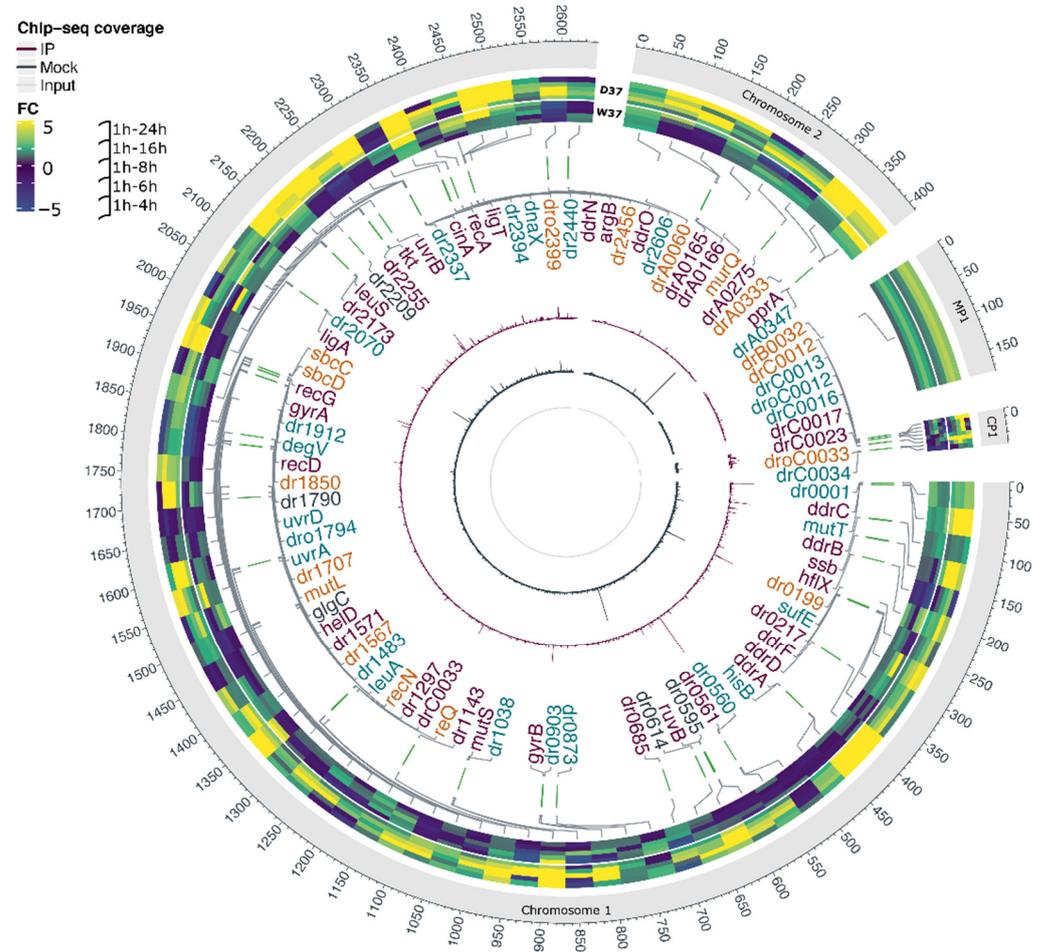

Figure 7. Representation of identified DdrO target genes across the *D. radiodurans* genome. Each genome replicon is represented by an outer circle. Heatmaps represent Fold Change values for the 85 genes, sometimes in operons, matching with at least two of the selected criteria (see Figure 1). The genomic positions of deregulated genes are drawn in grey connections. The DdrO bound sites associated with identified candidate genes included in the DdrO regulon are illustrated by green vertical lines. High confidence DdrO targets genes matching with three criteria are indicated in purple. The other genes matching with two criteria are labeled in dark gray (ChIP-seq and Transcriptomics), blue (Transcriptomics and RDRM), or orange (ChIP-seq and RDRM). The purple circle shows the mean coverage from the three IP replicates. The dark grey circle shows Input tag density profiles and the light grey shows the Mock tag density profiles.

Based on these results, the DdrO regulon comprises 16 previously predicted RDR genes, mainly involved in DNA repair pathways. Moreover, three other genes involved in DNA metabolism (*recG*, *helD*, and DNA ligase *ligA*), four genes associated with different metabolic pathways, five genes involved in translation, and seven new genes encoding proteins of unknown function (Figure 7, Table 2 and Table S11) are also under the control of DdrO. Surprisingly, two genes encoding transposases, *drC0017* and *dr1296*, also matched all the criteria. One copy of *drC0017* and two copies of the second IS (*dr1296* and *drC0033*) are present in the genome. According to the Chandler classification (https://www-is.biotoul.

fr/index.php) (accessed on 10 September 2021) *dr1296* encodes the IS*Dra5* and *drC0017* is part of the transposable element TnDra1. Because *drC0033* and *dr1296* CDS exhibited 100% sequence identity, we were not able to determine, by RNA-seq, whether one of the two genes displayed an upregulation during the time course. Moreover, only a small peak was observed for *drC0033* but was below the threshold established for this study.

In addition, 47 genes matched only two criteria. Among these, six genes were upregulated in D37 and a ChIP peak was located in each promoter region, but their sequences did not exhibit any RDRM or other conserved motif (Table S12). However, we cannot exclude the possibility that the DdrO protein bound to a degenerate RDRM sequence, which was not reported because of the criteria used here for in silico analyses. Among these, *dr2606* encodes a predicted primosomal protein N′ and *dr1790* encodes for the yellow protein, belonging to the ancient yellow/major royal jelly (MRJ) protein family. The deletion of *dr1790* in *D. radiodurans* increased its membrane permeability and decreased the cell growth rate and survival upon exposure to hydrogen peroxide and radiation [53].

Twenty-six other genes were only associated with a DdrO peak and also with an RDRM in their promotor region (Figure 7 and Table S13). From this set of 26 genes, 13 are divergently transcribed from genes that matched all criteria. It is thus likely that only one of the two genes that share the same intergenic region was regulated by DdrO. Seven of the 13 other genes, such as *dr0001* encoding DNA polymerase III subunit beta, were described as upregulated upon exposure to gamma rays [12], but was not reported as differentially expressed in a Δ*irrE* mutant [24]. The *uvrA* and *uvrD* genes were also reported as being upregulated in the first 1.5 h when cells are exposed to large doses of gamma rays [12] and were differentially expressed very early in a Δ*irrE* mutant. These results suggest that the expression of these genes, including *uvrA* and *uvrD*, may be augmented very early during the time course (<1 h) or are under the control of DdrO and other regulatory elements, and the simple depletion of DdrO did not modify their expression under our experimental procedures.

Finally, 16 genes were only found to be upregulated in D37 and not in W37, and an RDRM was detected near their respective promoter region, but no ChIP-Seq peak was reported from the ChIP-seq analysis (Table S14). However, careful inspection of all the peaks with the IGV program showed that a small peak, that fell below the threshold used to analyze ChIP-seq data, was observed in the promoter regions of five genes (*sbcD*, *recQ*, *drC0012* encoding a putative transcriptional regulator, *drC0033* (transposase), and *dr1707* encoding DNA polymerase I).

Table 2. List of genes belonging to the DdrO regulon in *D. radiodurans* matching all criteria (transcriptomic, ChIP-seq, and RDRM). Previous experiments: (1) RDR genes previously predicted by bioinformatic analyses [19,33] (2) by in vitro EMSA assays which analyzed the binding of DdrO to RDRM sequences located in the promoter regions of these genes [26] or shown to be regulated after exposure to radiation by (3) transcriptomic [12], (4) proteomic studies [54], or (5) in a Δ*irrE* mutant [24]. Logo of the RDRM consensus performed by MEME from all the RDRM sequences listed in Table 2. ND: Not determined.

ID (White et al. 1999) [48]	Local ID	Genes	Definition	Position Relative to Start Codon	RDRM	DdrO Dependent Induction Fold Change 1h/Time Point	Protein Induction with MMC/DdrO/IrrE Dependent	Predicted to RDR Regulon (Makarova et al. 2007; Blanchard et al. 2017) [19,33]	Previous Experiments
Replication, recombination, and repair									
DR0003	DRO_0003	*ddrC*	DdrC	−38	GTTATGTCAAAAACATAATC		ND	X	3, 5
DR0070	DRO_0070	*ddrB*	single-stranded DNA-binding protein	−31	TGTTATGTTATTTACGTAAG		ND	X	1, 2, 3, 4, 5
DR0100	DRO_0099	*ssb*	single-stranded DNA-binding protein	−118	TTTTATGTCATTGACATAAT		ND	X	1, 2, 3, 4, 5
DR0326	DRO_0323	*ddrD*	DNA repair protein	−29	ATTCTGCTAAAAACAGAATA		ND	X	1, 2, 5
DR0423	DRO_0421	*ddrA*	single-stranded DNA-binding protein	−23/−44	ATTCTGTTCTAAACTAAAT/ TTTATGTCTTGACCGTAAT		ND	X	1, 2, 3, 4, 5
DR0596	DRO_0596	*ruvB*	Holliday junction DNA helicase RuvB	−28	ATTTCGCAAATAGCGTAAT		ND	X	1, 2, 3, 5
DR0906	DRO_0899	*gyrB*	DNA topoisomerase IV subunit B	−157	ATTCTGTAAGAGACGTAAT		ND	X	1, 2, 3, 4, 5
DR1039	DRO_1033	*mutS*	DNA mismatch repair protein MutS	−43	GTTTCGCTCAGAACGTAAA		ND	X	1, 2
DR1916	DRO_1894	*recG*	ATP-dependent DNA helicase RecG	−17	GTTACGCTGTGAGCGAAAT		yes/yes		3
DR1572	DRO_1552	*helD*	DNA helicase	+9	TTTATGTCTCTGGCAGAAC		no/no		5

Table 2. *Cont.*

ID (White et al. 1999) [48]	Local ID	Genes	Definition	Position Relative to Start Codon	RDRM	DdrO Dependent Induction Fold Change 1h/Time Point	Protein Induction with MMC/DdrO/IrrE Dependent	Predicted to RDR Regulon (Makarova et al. 2007; Blanchard et al. 2017) [19,33]	Previous Experiments
DR1902	DRO_1880	*recD*	exodeoxyribonuclease V	−46	ATTACGCTGATGACATAAT		ND		5
DR1913	DRO_1891	*gyrA*	DNA gyrase subunit A	−118	ATTACGTGATTAACATAAT		ND	X	1, 2, 3, 4, 5
DR2069	DRO_2042	*ligA*	DNA ligase (NAD(+)) LigA	−47	ATTCTGCCCTGAACCGAAC		yes/yes		5
DR2275	DRO_2249	*uvrB*, putatively in operon with *dr2276*	excinuclease ABC subunit B	−105	CTTACGCTGTGGGCGTAAA		ND	X	1,2,3,5
DR2338	DRO_2308	*cinA*, first gene in operon with *ligT* and *recA*	Competence inducible protein A	−35	GTTATGCTGCTAGCAGAAA		ND	X	1,2,3,4,5
DRA0346	DRO_A0342	*pprA*	DNA repair protein	−29	AATCTGTTCAGGGCATAAT		ND	X	1,2,3,4,5
Regulation of transcription									
DR2574	DRO_2545	*ddrO*	transcriptional regulator	−153	ATTCTGTATTGACCGTAGC		ND	**X**	1,2,3,5

Table 2. *Cont.*

ID (White et al. 1999) [48]	Local ID	Genes	Definition	Position Relative to Start Codon	RDRM	DdrO Dependent Induction Fold Change 1h/Time Point	Protein Induction with MMC/DdrO/IrrE Dependent	Predicted to RDR Regulon (Makarova et al. 2007; Blanchard et al. 2017) [19,33]	Previous Experiments
Translation and post-translational modification									
DR0139	DRO_0139	h*flX*	GTPase HflX	−390	GTTCTGTCCGGGGCGAAAC		yes/no		4
DR2174	DRO_2145	*leuS*	leucine—tRNA ligase	−291	CATATGTCATGAGCATAAC		no/no		
DR2255	DRO_2229	Putatively n operon with *dr2254* d*drN*,	GNAT family N-acetyltransferase	−209/−193	AATACGCTAGGGGCGTAAA/ ATTCCGGTAAAGACAGAAT		no/no		3
DR2441	DRO_2415	putatively in operon with *dr2442*	acetyltransferase	−109	AATTTGTTATTTGCGAACT		yes/yes		3,5
Metabolism and metabolic transport									
DR0217	DRO_0217		thiosulfate sulfurtransferase	+4	ATTACGCCAAAGACGTGTT		no/no		4
DR0561	DRO_0559		sugar ABC transporter substrate-binding protein	−245	GTTCAGGAAAAAACATAAC		no protein observed		
DR1297	DRO_1288		ABC transporter	−131	GTTACGCTCCTAAACAAAT		no protein observed		
DR2256	DRO_2230	*tkt*	transketolase	−182/−199	TTTACGCCCCTAGCGTATT/ ATTCTGTCTTTACCGGAAT		ND	X	1,2,3
DRA0275	DRO_A0273		cytochrome C6	−93	ATGTGACAAAGAGCGTAT		no/no		5

Table 2. *Cont.*

ID (White et al. 1999) [48]	Local ID	Genes	Definition	Position Relative to Start Codon	RDRM	DdrO Dependent Induction Fold Change 1h/Time Point	Protein Induction with MMC/DdrO/IrrE Dependent	Predicted to RDR Regulon (Makarova et al. 2007; Blanchard et al. 2017) [19,33]	Previous Experiments
Unknown function									
DR0219	DRO_0219	*ddrF*	hypothetical protein	−43	TGTTATGTTATATACGTAAA		ND	X	1,2,3,5
DR0685	DRO_0681		hypothetical protein	−37	TCTTATGTTCTGAACGCTTT		no/no		
DR1143	DRO_1140		hypothetical protein	−47	GTTATGTTTTAAGCGTAAA		ND	X	1,2,3,5
DR1571	DRO_1551		ABC transporter	−290	GTTCTGCCAGAGACATAAA		no/no		
DR2173	DRO_2144		hypothetical protein	−17	GTTATGCTCATGACATATG		yes/yes		3,5
DRA0165	DRO_A0167	Putatively In operon with *drA0166*	Conserved hypothetical protein	−83	GGTTATGCTATTTACATAAC		yes/yes		5
DRC0023	DRO_C0021		hypothetical protein	−193	CTTTGTTCTGTTAGCCTAAC		no protein observed		3
ambiguous									
DRC0017	DRO_C0017	Putatively In operon with *drC0018*	transposase	−248	TAGTATGCTTCTGGCGTAGT		no protein observed		
DR1296/DRC0033	DRO_1287/ DRO_C0033		transposase	−52	AATATGTAAATAGCATAGT		no protein observed		
Consensus motif TATGYYHTTARCRKA									

A highly condensed structure of the *D. radiodurans* chromatin may have locally impaired or decreased the efficiency of the ChIP experiments, and thus, these genes may also belong to the DdrO regulon. The location of the RDRM was also analyzed in the promoter region of all genes matching two or three of our criteria. The RDRM sequences were mainly located in the vicinity of the predicted position of *E. coli*-like -35 and -10 promoter consensus sequences (Tables S9, S11, and S12). These results are consistent with previous reports showing that RDRM in *D. deserti* was found both upstream (-50 bp) and downstream ($+20$ bp) of transcriptional start sites (TSS) potentially overlapping with the RNA polymerase binding site [17].

In parallel, the amount of protein encoded by the newly identified genes was analyzed in wild-type cells and in a Δ*irrE* mutant after exposure to mitomycin C (MMC). For this purpose, we monitored, by Western blot analysis, the expression of C-terminal-tagged recombinant proteins. The cellular levels of five recombinant proteins (RecG, LigA, DdrN, DRA0166, and DR2173) increased in wild-type cells in response to MMC, but remained constant in an Δ*irrE* mutant, thus corroborating our data (Figure 8).

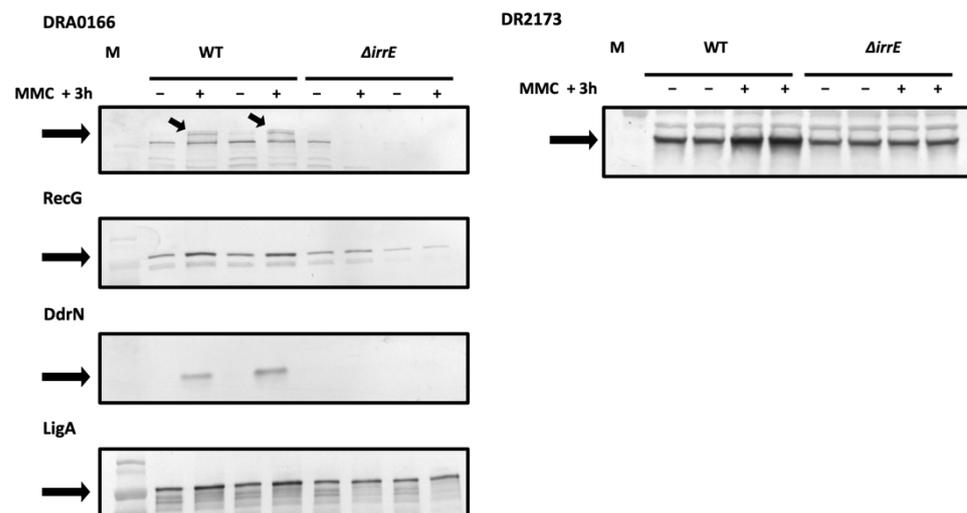**Figure 8.** Expression of several new DdrO target genes induced after exposure to MMC in an IrrE and DdrO-dependent manner. Δ*irrE* or wt cells expressing recombinant DRA0166-V5, RecG-V5, DdrN-V5, LigA-V5, and DR2173-V5 proteins (indicated by black arrows) were incubated (+) or not (−) with MMC (1 μg/mL) at 30 °C for 3 h. Cell extracts were subjected to SDS-PAGE and analyzed by Western blotting with anti-V5 antibodies. For DRA0165 and DR2173, 15 μg of proteins was loaded on each well, for RecG and DdrN, 10 μg of proteins, and for LigA, 5 μg.

We were not able to detect *dr1297* or *dr0561* tagged proteins, but these genes encode a predicted ABC transporter and a sugar ABC transporter, respectively, containing predicted transmembrane regions or a predicted periplasmic peptide signal that could lead to a low solubility of these proteins. In agreement with our gene expression data, previous transcriptome studies have shown that *dr0561* was not upregulated in a Δ*irrE* strain when compared to a wild-type strain upon exposure to γ-ray irradiation [24], lending support to the idea that this predicted transporter is regulated in a DdrO/IrrE manner.

Among the set of genes predicted in silico by Makarova et al. [33], the *frnE* and *rsr* genes were upregulated in D37 over the time course of the experiment, but no DdrO-peak was found in their respective promoter regions. We analyzed the expression of the FrnE and Rsr proteins, and of DdrR (DR0053) predicted as belonging to the RDR regulon in *D. deserti* [19], although *ddrR* was not upregulated during our time course experiment and no peak was observed in its promoter region. The cellular levels of three recombinant proteins increased in response to MMC in wild-type cells, and also in a Δ*irrE* mutant (Figure S6). These results strongly suggest that DdrR, FrnE, and Rsr proteins are induced by genotoxic stress but not in a DdrO/IrrE-dependent manner.

With the exception of HflX, which was found to be induced in wild-type and Δ*irrE* cells after exposure to MMC, no change in protein quantity was observed for DRA0275, DR0217, DR0685, DR1571, DR1572, DR2255, and DR2174 after exposure to MMC (Figure S6). However, as already reported in *D. radiodurans*, the upregulation of several genes at the transcriptomic level was not always observed at the proteomic level [54]. This may be due to their abundance in the cell, their very transient expression, or their instability. A mechanism of translational regulation may also occur after transcription of these genes. Alternatively, it is possible that the genotoxic conditions after MMC exposure did not exhibit appropriate deleterious effects in cells to trigger induction of these proteins when compared to an exposure to other stresses, such as gamma radiation, Methyl MethaneSulfonate (MMS), or desiccation [27].

4. Discussion

An RDR regulon was proposed several years ago based on the presence of the RDRM sequence, a common 17 bp palindromic sequence, located in the promoter region of the most highly ionizing radiation and desiccation upregulated genes in *D. radiodurans* and *D. geothermalis* [33]. To date, identification of putative DdrO target genes in *D. radiodurans* has been mostly proposed by a combination of bioinformatic analyses based on microarray gene expression data [33] and a validation, in vitro, by Electrophoretic Mobility Shift Assay (EMSA) experiments [26]. Here, we combined two large scale approaches to identify DdrO targets in vivo with reliable accuracy. Analysis among transcriptome data, identification of enriched DdrO binding sites, and the presence of an RDRM in the promoter region of *D. radiodurans* genes allowed us to identify at least 35 DdrO target genes matching all criteria (Table 2) and to propose other genes that may also be regulated by this transcription factor (Tables S10–S12).

Up to 70% of the identified target genes were previously predicted to be part of the RDR regulon [19,26,33]. The *ddrF* gene (*dr0219*), absent from the genome annotation published by Hua and Hua [49], is identified here as belonging to the RDR regulon, as initially described [33]. In addition, we highlighted 18 new DdrO target genes, including genes involved in DNA maintenance such as *dr1289*, *dr1572*, and *dr2069*, encoding the RecG helicase, HelD superfamily I helicase IV, and the replicative DNA ligase LigA, respectively. In agreement with these data, transcriptome analysis of cells recovering from exposure to ionizing radiation showed no upregulation of expression of these three genes in a Δ*irrE* mutant compared to the wild-type strain [24]. In *E. coli*, RecG plays an important role in DNA repair, recombination, and replication [55], whereas HelD, from *Bacillus subtilis* or *Mycobacterium smegmatis*, is associated with transcriptional pathways [56,57]. In *D. radiodurans*, cells devoid of RecG exhibit a delay in growth and double strand break (DSB) repair kinetics, and a decrease in resistance to γ-irradiation and H_2O_2 [58,59], whereas the Δ*dr1572* mutant exhibited a greater sensitivity to H_2O_2, but no change in resistance to ionizing radiation and to MMC when compared to the wild-type strain [60].

The identification of *ligA* as a target gene is interesting, because DNA ligases are implicated in DNA repair and are essential in other fundamental processes within the cell [61]. Ligase activity is crucial during DNA recombination and replication, explaining the constitutive expression of DNA ligase during all phases of the cell cycle [61]. Therefore, DdrO binding on the *ligA* promoter region should not completely repress gene expression, to thus ensure a minimum level of DNA ligase activity. However, in response to elevated amounts of DNA damage, and particularly to DSB, the basal level of LigA may not be sufficient for accurate Single Strand Annealing (SSA) [8] and Extended Synthesis-Dependent Strand Annealing (ESDSA) mechanisms, or for homologous recombination [5].

We identified new DdrO target genes, *ddrN* and *dr2255*, encoding putative GNAT family acetyltransferases that may be involved in post-translational modification (PTM) pathways. The Δ*irrE* cells recovering from exposure to ionizing radiation exhibited no upregulation of *ddrN* expression compared to a wild-type strain [24]. PTMs in bacteria play crucial roles in various cellular pathways, including after metabolic shifts and stress

adaptation [62]. Acetylation is known to modify a variety of substrates involved in RNA metabolism, enzymatic activity, or DNA-related mechanisms [63]. In *E. coli*, acetylation of the chromosomal replication initiation protein DnaA leads to an arrest of DNA replication [64]. Moreover, acetylation of histone-like nucleoid protein HU in *M. tuberculosis* alters the in vitro DNA-binding capacity of HU and the DNA structure, which may affect gene transcription and other protein–DNA interactions [65]. *D. radiodurans* HU protein has been reported as a major actor of nucleoid compaction [52,66] and may also be acetylated. Further analyses are required to understand the impact of acetylation activity in response to DNA damage in *D. radiodurans*.

Surprisingly, two genes, *drC0017* and *dr1287*, encoding transposases belonging to the Tn3 family and to the IS*Dra5* (IS*5* family), respectively, were identified as belonging to the DdrO regulon. Transposons are major actors of genome remodeling and play an important role to create diversity and to facilitate adaptation of the host to extreme environmental conditions. Insertion sequences are abundant in *D. radiodurans* and IS transposition is a major event in spontaneous, in addition to induced, mutagenesis [67]. It has been previously shown that IS*Dra2*, IS*Dra5*, IS*Dra3*, IS*Dra4*, and IS*2621* belonging to different families (IS*200*/IS*605*, IS*5*, IS*630*, IS*630*, and IS*4*, respectively) were transpositionally active in this organism under normal growth conditions and transposition was enhanced in cells recovering from DNA damage. Transposable element expression and movement are generally tightly regulated and different mechanisms control their gene expression. In *E. coli*, LexA protein represses expression of the Tn5 transposase gene [68]. Further studies are required to better understand how DdrO contributes to the regulation of transposition events of these two ISs.

We also identified several genes encoding proteins of unknown function, such as *dr2173* and the *drA0165-drA0166* operon, which were strongly upregulated in response to a depletion of DdrO. These genes were not upregulated in irradiated Δ*irrE* cells [24]. Domains of unknown function (DUF) found in DR2173 and DRA0166 are widely conserved in bacteria. A DUF4132 within DRA0166 may be involved in the molybdopterin biosynthesis. DR2173 also shares an N-terminal WGR domain with the MolR protein, which may be involved in regulation of molybdate biosynthesis in *E. coli* [69] and was described as belonging to the LexA regulon. However, molybdate-metabolism associated genes, such as *D. radiodurans moeA* or *moeB*, were not differentially expressed in D37 compared to W37. The role of these strongly upregulated unknown genes in response to DNA damage also remains to be discovered.

DdrO-structure and biochemicals data suggested that, in response to DNA damaging conditions, upregulation of the expression of genes of the RDR regulon is dependent on a dynamic balance between DdrO dimers bound to DNA and the IrrE-cleavable DdrO monomer forms [25,28]. Thus, the cleavage of DdrO monomers by IrrE would reduce the amount of DdrO dimers able to bind to DNA, leading to the induction of the expression of genes controlled by this regulator. Here, the transcriptome data showed that identified RDR regulon genes were not all upregulated at the same time during the DdrO depletion. The expression of some genes, such as *ddrC*, *ddrD*, and *pprA*, was strongly upregulated at early times, whereas *ssb*, *uvrD*, and *ddrF* showed a late upregulation, suggesting that DdrO bound with more or less affinity to DNA according to the divergent RDRM sequences that may diverge. Therefore, after exposure of ionizing radiation, some genes would be upregulated earlier than others during cell recovery. It has been shown that, following extended DdrO-depletion, *D. radiodurans* cells were engaged in an apoptotic-like cell death (ALD) pathway leading to morphological alterations, such as larger cells, membrane blebbing, and DNA fragmentation [20]. In *Caulobacter crescentus*, the BapE endonuclease was reported to be involved in DNA fragmentation upon severe and extensive DNA damage [70]. BapE induction is triggered only in the case of prolonged LexA self-cleavage and was not described as part of early induced SOS response genes [70]. Instead, our results did not allow us to identify new genes belonging to the RDR regulon and differentially expressed at late times, suggesting that ALD would be triggered by long-lasting induction

of one or more genes from the RDR regulon, or by a cascade of regulation events following depletion of DdrO.

Despite the identification of many promoter regions containing a putative RDRM sequence in the *D. radiodurans* genome, we showed that only a small proportion of these are bound by DdrO. Based on the presence of the RDRM sequence and induction of expression in response to ionizing radiation, *rsr*, *frnE*, *irrI*, *ddrR*, and the *hutU* operon were previously described as part of the RDR regulon in *D. radiodurans* [19,33]. In agreement with Wang et al. [26], we showed that expression of the *hutU* operon, *rsR*, *frnE*, and *irrI* is not under the control of DdrO, but also *ddrR*. However, we showed that the quantity of RsR, FrnE, and DdrR proteins is induced in response to exposure to MMC (Figure S6) in concordance with gene expression data [12,13,19]. Thus, identification of the RDRM sequence is not sufficient to enable DdrO binding. Many factors, such as DNA structure or binding competition between multiple transcription factors, can affect accessibility to a genome region for DdrO [71]. If our data showed that DdrO appears to bind exclusively to the RDRM sequence, then some RDR regulon genes, such as *ddrA*, *gyrA*, or *dr1572*, were also described as being under the control of another major regulator, DdrI [72], highlighting the regulatory cross-talk in the *D. radiodurans* DNA damage response.

The *D. radiodurans* genome encodes more than one hundred predicted transcriptional regulators, but few studies have been undertaken to identify the genes they may regulate. The RDR regulon of the radioresistant bacterium *D. radiodurans*, characterized here by integrative genomic analyses, paves the way for further studies to better depict the regulatory networks underlying the mechanisms that contribute to the extreme radiation tolerance of this fascinating bacterium.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cells10102536/s1. Figure S1: A *ddrO* allele replacement by a gene expressing the recombinant DdrO-V5 protein did not modify the growth rate. Figure S2. Visualization, through IGV, of the tag density profiles for 20 genes reported as belonging to the RDR regulon. Figure S3. Western blot analysis of the expression of the recombinant RecA-HA, PprA-HA, and DdrD-HA proteins in D37 at 37 °C. Figure S4. Visualization using Tablet software [73] of the DdrO gene reads mapping at 4 h at the *D. radiodurans ddrO* genome locus and on the replication plasmids in strains W37 and D37, respectively. Figure S5. Hierarchical clustering of genes whose expression is specifically downregulated in response to DdrO depletion. Figure S6. Expression of several proteins after exposure to MMC in an IrrE- and DdrO-independent manner. Table S1. Bacterial strains and plasmids used in this study. Table S2. Primers used for strain and plasmid constructions. Restriction sites are indicated in bold. Table S3. List of CDS in the *D. radiodurans* strain ATCC 13939 genome sequence. The paralogs found in this sequence, and the orthologs in other *D. radiodurans* R1 sequence releases, are indicated with a threshold of 80% of the maximum bit score applied. Table S4. Data integration for all CDS. List of CDS (Column A). RDR genes predicted by Makarova et al. [33] (Column B). RNA-seq data (columns C-F): Number of time points with differential expression in W37 (column C) or in D37 (column E), or only upregulated at intermediate time points, i.e., 6, 8, and 16 h in W37 (column D) or in D37 (column F). ChIP-seq data (Columns G-L). Number of peaks detected in the promoter region of each CDS (Column G), peak name (Column H), and their coordinates on Genome Orsay (Columns K-L). In the case of the presence of two peaks, all of the information for the second peak is given in columns Y-AF. Search for palindromic or non-palindromic motifs ("palindrome" or "no palindrome", respectively) with MEME or/and FIMO with an occurrence of one motif per sequence or any number of repetitions ("0–1" or "any", respectively) (Columns M-X). The sequences of each motif close to the RDRM is indicated for each peak. For palindromic motifs, only the sequence found on one DNA strand is indicated. Table S5. Conserved motifs predicted by FIMO. The four position specific scoring matrix (PSSM) from MEME analysis (Figure 1B) was used as inputs for the FIMO program, scanning sequences between −500 and +100 of all annotated CDS, for palindromic or non-palindromic motifs with an occurrence of one motif per sequence or any number of repetitions. For a search for palindromic sequences, FIMO listed the motifs found on the both strands. Table S6. Number of unique and common up- and downregulated genes in W37 and D37, respectively, during the time course. For each time lapse, a 2-fold change expression threshold for the ratio experiments was applied together with a *p*-value < 0.01. Table S7. Variation

of all CDS expression during the time course of W37 and D37 cells, in addition to *ddrO* and *spr* genes cloned in the *repU$_{Ts}$* or the *repU$^+$* replication vectors, comparing each time point of the kinetic (4, 6, 8, 16, and 24 h) to the first (1 h). Red denotes upregulated genes (FC $\geq$ 2, *p*-value $\leq$ 0.01) and green denotes downregulated genes (FC $\leq$ −2, *p*-value $\leq$ 0.01). Genes that were previously predicted by Makarova [33] are underlined in yellow. Tables S8–S10. List of all the deregulated genes (Table S8), upregulated (Table S9) or downregulated (Table S10) as differentially expressed in the W37 strain during the time periods 1–6 h and 1–16 h in two or fewer comparisons (DE $\leq$ 2), and in more than three comparisons in the D37 strain (DE > 3). Genes that were previously predicted by Makarova et al. [33] are underlined in yellow. Tables S11–S14. List of genes belonging to the RDR regulon in *D. radiodurans* (Table S11) or matching two criteria only i.e., 6 genes with ChIP-seq and Transcriptomics criteria (Table S12), 16 genes with Transcriptomics and RDRM criteria (Table S13), and 25 genes with ChIP-seq and RDRM criteria (Table S14) (Lists L1–L4, Figure 1D). The coordinates of the ChIP-seq peaks and the genes regulated by DdrO are indicated for each replicon (Dro on Chr I, Dro_A: Chr II, Dro_C: plasmid). For RNA-seq assays, gene expression ratios in D37 and W37 species are reported for each time lapse using time 1 h as a reference. Red denotes upregulated genes (FC $\geq$ 2, *p*-value $\leq$ 0.01) and green denotes downregulated genes (FC $\leq$ −2, *p*-value $\leq$ 0.01). The predicted position of *E. coli*-like −35 and −10 promoter consensus sequences was carried out by BPROM [47]; in addition, the sequence and the position of the RDRM are highlighted from the predicted start of translation. Previous experiments: (1) RDR genes previously predicted by bioinformatic analyses [33], (2) by in vitro EMSA assays which analyzed the binding of DdrO to RDRM sequences located in the promoter regions of these genes [26] or shown to be regulated after exposure to radiation by (3) transcriptomic [12], (4) proteomic studies [54], or (5) in a Δ*irrE* mutant [24], respectively. Genes that were previously predicted by Makarova et al. [33] are also underlined in yellow. File S1. Script used in our studies to define without a priori different lists of candidate genes to be DdrO targets.

Author Contributions: Conceptualization, F.C. and P.S.; Methodology, F.C., G.L. and Y.Z.; Software, F.C. and G.L.; Validation, N.E., F.C. and G.L.; Formal Analysis, N.E., F.C., Y.Z. and G.L.; Investigation, N.E., C.B.d.l.T.; G.C., E.B. and P.S.; Resources, P.S.; Data Curation, N.E., F.C., G.L. and Y.Z.; Writing—Original Draft Preparation, F.C. and N.E.; Writing—Review and Editing, N.E., C.B.d.l.T., E.B., P.S. and F.C.; Visualization, N.E. and F.C.; Supervision, F.C. and P.S.; Project Administration, F.C. and P.S.; Funding Acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: ID of genes and proteins described in this study as well as their predicted function are available in Table 2 and Supplementary Tables S7–S14.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Blasius, M.; Hubscher, U.; Sommer, S. Deinococcus radiodurans: What belongs to the survival kit? *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 221–238. [CrossRef]
2. Confalonieri, F.; Sommer, S. Bacterial and archael resistance to ionizing radiation. *J. Phys.* **2011**, *261*, 012005. [CrossRef]
3. Daly, M.J. Death by protein damage in irradiated cells. *DNA Repair* **2012**, *11*, 12–21. [CrossRef] [PubMed]
4. Slade, D.; Radman, M. Oxidative stress resistance in Deinococcus radiodurans. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 133–191. [CrossRef]
5. Zahradka, K.; Slade, D.; Bailone, A.; Sommer, S.; Averbeck, D.; Petranovic, M.; Lindner, A.B.; Radman, M. Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature* **2006**, *443*, 569–573. [CrossRef]

6. Bouthier de la Tour, C.; Boisnard, S.; Norais, C.; Toueille, M.; Bentchikou, E.; Vannier, F.; Cox, M.M.; Sommer, S.; Servant, P. The deinococcal DdrB protein is involved in an early step of DNA double strand break repair and in plasmid transformation through its single-strand annealing activity. *DNA Repair* **2011**, *10*, 1223–1231. [CrossRef] [PubMed]

7. Slade, D.; Lindner, A.B.; Paul, G.; Radman, M. Recombination and replication in DNA repair of heavily irradiated *Deinococcus radiodurans*. *Cell* **2009**, *136*, 1044–1055. [CrossRef]

8. Xu, G.; Lu, H.; Wang, L.; Chen, H.; Xu, Z.; Hu, Y.; Tian, B.; Hua, Y. DdrB stimulates single-stranded DNA annealing and facilitates RecA-independent DNA repair in *Deinococcus radiodurans*. *DNA Repair* **2010**, *9*, 805–812. [CrossRef]

9. Floc'h, K.; Lacroix, F.; Servant, P.; Wong, Y.S.; Kleman, J.P.; Bourgeois, D.; Timmins, J. Cell morphology and nucleoid dynamics in dividing *Deinococcus radiodurans*. *Nat. Commun.* **2019**, *10*, 3815. [CrossRef] [PubMed]

10. Zimmerman, J.M.; Battista, J.R. A ring-like nucleoid is not necessary for radioresistance in the Deinococcaceae. *BMC Microbiol.* **2005**, *5*, 17. [CrossRef]

11. 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. Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *PLoS ONE* **2010**, *5*, e12570. [CrossRef]

12. Liu, Y.; Zhou, J.; Omelchenko, M.V.; Beliaev, A.S.; Venkateswaran, A.; Stair, J.; Wu, L.; Thompson, D.K.; Xu, D.; Rogozin, I.B.; et al. Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4191–4196. [CrossRef]

13. Tanaka, M.; Earl, A.M.; Howell, H.A.; Park, M.J.; Eisen, J.A.; Peterson, S.N.; Battista, J.R. Analysis of *Deinococcus radiodurans*'s transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics* **2004**, *168*, 21–33. [CrossRef] [PubMed]

14. Kreuzer, K.N. DNA damage responses in prokaryotes: Regulating gene expression, modulating growth patterns, and manipulating replication forks. *Cold Spring Harb. Perspect Biol.* **2013**, *5*, a012674. [CrossRef]

15. Bonacossa de Almeida, C.; Coste, G.; Sommer, S.; Bailone, A. Quantification of RecA protein in *Deinococcus radiodurans* reveals involvement of RecA, but not LexA, in its regulation. *Mol. Genet. Genom.* **2002**, *268*, 28–41. [CrossRef] [PubMed]

16. Narumi, I.; Satoh, K.; Kikuchi, M.; Funayama, T.; Yanagisawa, T.; Kobayashi, Y.; Watanabe, H.; Yamamoto, K. The LexA protein from *Deinococcus radiodurans* is not involved in RecA induction following gamma irradiation. *J. Bacteriol.* **2001**, *183*, 6951–6956. [CrossRef] [PubMed]

17. Ludanyi, M.; Blanchard, L.; Dulermo, R.; Brandelet, G.; Bellanger, L.; Pignol, D.; Lemaire, D.; de Groot, A. Radiation response in *Deinococcus deserti*: IrrE is a metalloprotease that cleaves repressor protein DdrO. *Mol. Microbiol.* **2014**, *94*, 434–449. [CrossRef] [PubMed]

18. Vujicic-Zagar, A.; Dulermo, R.; Le Gorrec, M.; Vannier, F.; Servant, P.; Sommer, S.; de Groot, A.; Serre, L. Crystal structure of the IrrE protein, a central regulator of DNA damage repair in deinococcaceae. *J. Mol. Biol.* **2009**, *386*, 704–716. [CrossRef]

19. Blanchard, L.; Guerin, P.; Roche, D.; Cruveiller, S.; Pignol, D.; Vallenet, D.; Armengaud, J.; de Groot, A. Conservation and diversity of the IrrE/DdrO-controlled radiation response in radiation-resistant Deinococcus bacteria. *Microbiologyopen* **2017**, *6*, e00477. [CrossRef]

20. Devigne, A.; Ithurbide, S.; Bouthier de la Tour, C.; Passot, F.; Mathieu, M.; Sommer, S.; Servant, P. DdrO is an essential protein that regulates the radiation desiccation response and the apoptotic-like cell death in the radioresistant *Deinococcus radiodurans* bacterium. *Mol. Microbiol.* **2015**, *96*, 1069–1084. [CrossRef]

21. Bouthier de la Tour, C.; Mathieu, M.; Meyer, L.; Dupaigne, P.; Passot, F.; Servant, P.; Sommer, S.; Le Cam, E.; Confalonieri, F. In vivo and in vitro characterization of DdrC, a DNA damage response protein in *Deinococcus radiodurans* bacterium. *PLoS ONE* **2017**, *12*, e0177751. [CrossRef] [PubMed]

22. Earl, A.M.; Mohundro, M.M.; Mian, I.S.; Battista, J.R. The IrrE protein of *Deinococcus radiodurans* R1 is a novel regulator of *recA* expression. *J. Bacteriol.* **2002**, *18*, 6216–6224. [CrossRef] [PubMed]

23. Hua, Y.; Narumi, I.; Gao, G.; Tian, B.; Satoh, K.; Kitayama, S.; Shen, B. PprI: A general switch responsible for extreme radioresistance of *Deinococcus radiodurans*. *Biochem. Biophys. Res. Commun.* **2003**, *306*, 354–360. [CrossRef]

24. Lu, H.; Chen, H.; Xu, G.; Shah, A.M.; Hua, Y. DNA binding is essential for PprI function in response to radiation damage in *Deinococcus radiodurans*. *DNA Repair* **2012**, *11*, 139–145. [CrossRef]

25. Magerand, R.; Rey, P.; Blanchard, L.; de Groot, A. Redox signaling through zinc activates the radiation response in *Deinococcus* bacteria. *Sci. Rep.* **2021**, *11*, 4528. [CrossRef] [PubMed]

26. Wang, Y.; Xu, Q.; Lu, H.; Lin, L.; Wang, L.; Xu, H.; Cui, X.; Zhang, H.; Li, T.; Hua, Y. Protease activity of PprI facilitates DNA damage response: Mn^{2+}-dependence and substrate sequence-specificity of the proteolytic reaction. *PLoS ONE* **2015**, *10*, e0122071. [CrossRef]

27. Narasimha, A.; Basu, B. New insights into the activation of radiation desiccation response regulon in *Deinococcus radiodurans*. *J. Biosci.* **2021**, *46*, 10. [CrossRef]

28. de Groot, A.; Siponen, M.I.; Magerand, R.; Eugenie, N.; Martin-Arevalillo, R.; Doloy, J.; Lemaire, D.; Brandelet, G.; Parcy, F.; Dumas, R.; et al. Crystal structure of the transcriptional repressor DdrO: Insight into the metalloprotease/repressor-controlled radiation response in Deinococcus. *Nucleic. Acids. Res.* **2019**, *47*, 11403–11417. [CrossRef]

29. Lu, H.; Wang, L.; Li, S.; Pan, C.; Cheng, K.; Luo, Y.; Xu, H.; Tian, B.; Zhao, Y.; Hua, Y. Structure and DNA damage-dependent derepression mechanism for the XRE family member DG-DdrO. *Nucleic. Acids. Res.* **2019**, *47*, 9925–9933. [CrossRef] [PubMed]

30. Lim, S.; Jung, J.H.; Blanchard, L.; de Groot, A. Conservation and diversity of radiation and oxidative stress resistance mechanisms in Deinococcus species. *FEMS Microbiol. Rev.* **2019**, *43*, 19–52. [CrossRef] [PubMed]

31. Blanchard, L.; de Groot, A. Coexistence of SOS-dependent and SOS-independent regulation of DNA repair genes in radiation-resistant Deinococcus bacteria. *Cells* **2021**, *10*, 924. [CrossRef] [PubMed]

32. de Groot, A.; Dulermo, R.; Ortet, P.; Blanchard, L.; Guerin, P.; Fernandez, B.; Vacherie, B.; Dossat, C.; Jolivet, E.; Siguier, P.; et al. Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium *Deinococcus deserti*. *PLoS Genet.* **2009**, *5*, e1000434. [CrossRef] [PubMed]

33. Makarova, K.S.; Omelchenko, M.V.; Gaidamakova, E.K.; Matrosova, V.Y.; Vasilenko, A.; Zhai, M.; Lapidus, A.; Copeland, A.; Kim, E.; Land, M.; et al. Deinococcus geothermalis: The pool of extreme radiation resistance genes shrinks. *PLoS ONE* **2007**, *2*, e955. [CrossRef] [PubMed]

34. Anaganti, N.; Basu, B.; Mukhopadhyaya, R.; Apte, S.K. Proximity of radiation desiccation response motif to the core promoter is essential for basal repression as well as gamma radiation-induced *gyrB* gene expression in *Deinococcus radiodurans*. *Gene* **2017**, *615*, 8–17. [CrossRef]

35. Meima, R.; Lidstrom, M.E. Characterization of the minimal replicon of a cryptic *Deinococcus radiodurans* SARK plasmid and development of versatile *Escherichia coli-D. radiodurans* shuttle vectors. *Appl. Environ. Microbiol.* **2000**, *66*, 3856–3867. [CrossRef]

36. Mennecier, S.; Coste, G.; Servant, P.; Bailone, A.; Sommer, S. Mismatch repair ensures fidelity of replication and recombination in the radioresistant organism *Deinococcus radiodurans*. *Mol. Genet. Genom.* **2004**, *272*, 460–469. [CrossRef]

37. Dardalhon-Samsonoff, M.; Averbeck, D. DNA-membrane complex restoration in Micrococcus radiodurans after X-irradiation: Relation to repair, DNA synthesis and DNA degradation. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **1980**, *38*, 31–52. [CrossRef] [PubMed]

38. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef]

39. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome. Res.* **2017**, *27*, 722–736. [CrossRef]

40. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Subgroup Genome Project Data Processing. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]

41. Danecek, P.; McCarthy, S.A. BCFtools/csq: Haplotype-aware variant consequences. *Bioinformatics* **2017**, *33*, 2037–2039. [CrossRef]

42. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [CrossRef]

43. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef] [PubMed]

44. Hyatt, D.; LoCascio, P.F.; Hauser, L.J.; Uberbacher, E.C. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* **2012**, *28*, 2223–2230. [CrossRef] [PubMed]

45. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef]

46. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome. Biol.* **2010**, *11*, R106. [CrossRef]

47. Solovyev, V.; Salamov, A.; Seledtsov, I.; Vorobyev, D.; Bachinsky, A. Automatic annotation of bacterial community sequences and application to infections diagnostic. In Proceedings of the International Conference on Bioinformatics Models, Rome, Italy, 26–29 January 2011; pp. 346–353. [CrossRef]

48. 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.; et al. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* **1999**, *286*, 1571–1577. [CrossRef] [PubMed]

49. Hua, X.; Hua, Y. Improved complete genome sequence of the extremely radioresistant bacterium *Deinococcus radiodurans* R1 obtained using PacBio single-molecule sequencing. *Genome Announc.* **2016**, *4*, e00886-16. [CrossRef]

50. Merhej, J.; Frigo, A.; Le Crom, S.; Camadro, J.M.; Devaux, F.; Lelandais, G. bPeaks: A bioinformatics tool to detect transcription factor binding sites from ChIPseq data in yeasts and other organisms with small genomes. *Yeast* **2014**, *31*, 375–391. [CrossRef]

51. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic. Acids. Res.* **2015**, *43*, W39–W49. [CrossRef] [PubMed]

52. Nguyen, H.H.; Bouthier de la Tour, C.; Toueille, M.; Vannier, F.; Sommer, S.; Servant, P. The essential histone-like protein HU plays a major role in *Deinococcus radiodurans* nucleoid compaction. *Mol. Microbiol.* **2009**, *73*, 240–252. [CrossRef] [PubMed]

53. Cheng, J.; Wang, H.; Xu, X.; Wang, L.; Tian, B.; Hua, Y. Characteristics of *dr1790* disruptant and its functional analysis in *Deinococcus radiodurans*. *Braz. J. Microbiol.* **2015**, *46*, 601–611. [CrossRef]

54. Basu, B.; Apte, S.K. Gamma radiation-induced proteome of *Deinococcus radiodurans* primarily targets DNA repair and oxidative stress alleviation. *Mol. Cell Proteom.* **2012**, *11*, M111.011734. [CrossRef] [PubMed]

55. Lloyd, R.G.; Rudolph, C.J. 25 years on and no end in sight: A perspective on the role of RecG protein. *Curr. Genet.* **2016**, *62*, 827–840. [CrossRef] [PubMed]

56. Kaur, G.; Kapoor, S.; Thakur, K.G. *Bacillus subtilis* HelD, an RNA polymerase interacting helicase, forms amyloid-like fibrils. *Front Microbiol.* **2018**, *9*, 1934. [CrossRef]

57. Newing, T.P.; Oakley, A.J.; Miller, M.; Dawson, C.J.; Brown, S.H.J.; Bouwer, J.C.; Tolun, G.; Lewis, P.J. Molecular basis for RNA polymerase-dependent transcription complex recycling by the helicase-like motor protein HelD. *Nat. Commun.* **2020**, *11*, 6420. [CrossRef] [PubMed]

58. Wu, Y.; Chen, W.; Zhao, Y.; Xu, H.; Hua, Y. Involvement of RecG in H2O2-induced damage repair in *Deinococcus radiodurans*. *Can. J. Microbiol.* **2009**, *55*, 841–848. [CrossRef] [PubMed]

59. Jeong, S.W.; Kim, M.K.; Zhao, L.; Yang, S.K.; Jung, J.H.; Lim, H.M.; Lim, S. Effects of conserved wedge domain residues on DNA binding activity of *Deinococcus radiodurans* RecG helicase. *Front Genet.* **2021**, *12*, 634615. [CrossRef]

60. Cao, Z.; Julin, D.A. Characterization in vitro and in vivo of the DNA helicase encoded by *Deinococcus radiodurans* locus DR1572. *DNA Repair* **2009**, *8*, 612–619. [CrossRef] [PubMed]

61. Wilkinson, A.; Day, J.; Bowater, R. Bacterial DNA ligases. *Mol. Microbiol.* **2001**, *40*, 1241–1248. [CrossRef]

62. Macek, B.; Forchhammer, K.; Hardouin, J.; Weber-Ban, E.; Grangeasse, C.; Mijakovic, I. Protein post-translational modifications in bacteria. *Nat. Rev. Microbiol.* **2019**, *17*, 651–664. [CrossRef]

63. Carabetta, V.J.; Cristea, I.M. Regulation, function, and detection of protein acetylation in bacteria. *J. Bacteriol.* **2017**, *199*, e00107-17. [CrossRef]

64. Zhang, Q.; Zhou, A.; Li, S.; Ni, J.; Tao, J.; Lu, J.; Wan, B.; Li, S.; Zhang, J.; Zhao, S.; et al. Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in *Escherichia coli*. *Sci. Rep.* **2016**, *6*, 30837. [CrossRef]

65. Ghosh, S.; Padmanabhan, B.; Anand, C.; Nagaraja, V. Lysine acetylation of the *Mycobacterium tuberculosis* HU protein modulates its DNA binding and genome organization. *Mol. Microbiol.* **2016**, *100*, 577–588. [CrossRef]

66. Toueille, M.; Mirabella, B.; Guerin, P.; Bouthier de la Tour, C.; Boisnard, S.; Nguyen, H.H.; Blanchard, L.; Servant, P.; de Groot, A.; Sommer, S.; et al. A comparative proteomic approach to better define Deinococcus nucleoid specificities. *J. Proteom.* **2012**, *75*, 2588–2600. [CrossRef]

67. Mennecier, S.; Servant, P.; Coste, G.; Bailone, A.; Sommer, S. Mutagenesis via IS transposition in *Deinococcus radiodurans*. *Mol. Microbiol.* **2006**, *59*, 317–325. [CrossRef]

68. Kuan, C.T.; Tessman, I. LexA protein of *Escherichia coli* represses expression of the Tn5 transposase gene. *J. Bacteriol.* **1991**, *173*, 6406–6410. [CrossRef] [PubMed]

69. Lee, J.H.; Wendt, J.C.; Shanmugam, K.T. Identification of a new gene, *molR*, essential for utilization of molybdate by *Escherichia coli*. *J. Bacteriol.* **1990**, *172*, 2079–2087. [CrossRef] [PubMed]

70. Bos, J.; Yakhnina, A.A.; Gitai, Z. BapE DNA endonuclease induces an apoptotic-like response to DNA damage in Caulobacter. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18096–18101. [CrossRef] [PubMed]

71. Gomes, A.L.; Wang, H.H. The role of genome accessibility in transcription factor binding in bacteria. *PLoS Comput. Biol.* **2016**, *12*, e1004891. [CrossRef]

72. Meyer, L.; Coste, G.; Sommer, S.; Oberto, J.; Confalonieri, F.; Servant, P.; Pasternak, C. DdrI, a cAMP receptor protein family member, acts as a major regulator for adaptation of *Deinococcus radiodurans* to various stresses. *J. Bacteriol.* **2018**, *200*, e00129-18. [CrossRef] [PubMed]

73. Milne, I.; Stephen, G.; Bayer, M.; Cock, P.J.; Pritchard, L.; Cardle, L.; Shaw, P.D.; Marshall, D. Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform.* **2013**, *14*, 193–202. [CrossRef] [PubMed]