Supplementary information for:

Probing the sORF-encoded peptides of *Deinococcus radiodurans* in response to extreme stress

Congli Zhou¹,*, Qianqian Wang¹,*, Yin Huang¹,*, Zijing Chen², Shuo Chen¹, Ye Zhao²,³,* and Chenxi Jia¹,*

¹ State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Lifeomics, National Center for Protein Sciences (The PHOENIX Center, Beijing), Beijing, China
² Institute of Biophysics, College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang, China.
³ Cancer Center, Zhejiang University, Hangzhou, Zhejiang, China.

# These authors contributed equally to this work.
* Corresponding author

Correspondence: Prof. Chenxi Jia and Prof. Ye Zhao

E-mail: cjia@ncpsb.org.cn (C.J.), yezhao@zju.edu.cn (Y.Z.)

Lead contact: Prof. Chenxi Jia,

Key words: small open reading frames (sORFs); sORF-encoded peptides (SEPs); peptidomics; oxidative stress; *Deinococcus radiodurans*
### Supplemental Table S1-1. Strains and plasmids used in this study.

| Strain and plasmid | Description | Source |
|--------------------|-------------|--------|
| **Strains**        |             |        |
| *E. coli*          |             |        |
| Trans5α            | Commercial  | TransGene |
| *D. radiodurans*   |             |        |
| R1                 | Wild type strain ATCC13939 | Lab stock |
| ΔSEP068184         | SEP068184 knockout strain, Str<sup>R</sup> | This study |
| ΔSEP068184-Cwt     | SEP068184 compensatory strain, Str<sup>R</sup>+Chl<sup>R</sup> | This study |
| **Plasmids**       |             |        |
| pRAD-3His-SEP068184| pRAD-N-3 × His containing wild type SEP068184 | This study |
| pRAD-SEP068184-3His | pRAD-C-3 × His containing wild type SEP068184 | This study |
| pRAD-3flag-SEP068184| pRAD-N-3 × flag containing wild type SEP068184 | This study |
| pRAD-SEP068184-3flag| pRAD-C-3 × flag containing wild type SEP068184 | This study |
| pRAD-SEP068184     | pRAD containing wild type SEP068184 gene, Amp<sup>R</sup> | This study |

### Supplemental Table S1-2. Primers or oligonucleotides used in this study.

| Name             | Sequence (5’→3’)               |
|------------------|--------------------------------|
| SEP068184 P1     | CACACCCGCAGTGGGGAAG             |
| SEP068184 P2     | CCAAGCTTTGCGACGCTGGGCTGGGCTG   |
| SEP068184 P3     | CGCGGATCCGACGACGACGGTGAAAC     |
| SEP068184 P4     | ATATAGACTTCCCTGGGCTGCGATCTGAC  |
| SEP068184 P5     | GCTCAGGAGGACCATATGGTGCTGGGCTTGACTGACC    |
| PRAD-F           | CCTGCAGGTCGAGTCGATCCTACAGGCCCCCAGCACG |
| PRAD-R           | GCTGGCGGCGGCCATATGCTGAGATGATGGCGGCCAGATCCGATT |
| C-6HIS-SEP068184-F | AATCGGATCCCTAGTGGATGATGATGATGATGGCGGCGCAG |
| C-6HIS-SEP068184-R | GGAGGATCCCATATGCGCAGGATGATGATGATGGCGGCGCAG |
| N-6HIS-SEP068184-F | CTG |
| N-6HIS-SEP068184-R | CAGTACCCGCACGTCGGATGATGATGATGATGATGCTATATGGGTC |
| SEP068184-R      | CTCC |
| Name           | Sequence (5’→3’)                                                                 |
|---------------|---------------------------------------------------------------------------------|
| STR-F         | ACGCGGATCCCTAGAAAAACTCATCGAGC                                                   |
| STR-R         | ATAACCATGGCTGCAAGACGCGCCTGCT                                                    |
| A-NEGFP-F     | GAGCTGTACAAGGACATATGGTGTGCGGTTACTGACCG                                          |
| CEGFP-068184-R| CTGCAGCTTGCTCACCATATGGCCCTCCCCGCCAGCGCGGTT                                      |
| A-UP-F        | GAGTGACATCTGACTGCTGCTGCS                                                        |
| A-DOWN-F      | GAGTGAGGGTTTTACTCAGCGC                                                          |
| B-UP-F        | CAGAAAAGGACGGGTAGCGAG                                                          |
| B-DOWN-F      | CGAAACTTACCTCCTCATCCTGCC                                                        |
| C-UP-F        | CGTGCGTGCAGTTATCTC                                                              |
| C-DOWN-F      | GAGCGTTTACGAAACCCAG                                                            |
| GROEL-F       | GAAATCAAGAAGCTGGCCGTG                                                           |
| GROEL-R       | CACCGGGACATGTCTTTC                                                            |
| A-UP-2R       | GCTTGTCTGAAAGGTCTGCTG                                                          |
| A-DOWN-2R     | CTCGCGGCTGTTCTCTGTCG                                                           |
| B-UP-2R       | CATGTCGTAAGGGCGGTACG                                                           |
| B-DOWN-2R     | CTCGGCGTGCCTTGAGGTGAGG                                                         |
| C-UP-2R       | CCCAGCAGCGAGTCTG                                                              |
| C-DOWN-2R     | CTGGCGGCTGAGGCGGAC                                                            |
| GROEL-2R      | CGGCGTCTGAGGCGGAC                                                             |
| A-UP-2R       | GCTTGTCTGAAAGGTCTGCTG                                                          |
| A-DOWN-2R     | CTCGCGGCTGTTCTCTGTCG                                                           |
| B-UP-2R       | CATGTCGTAAGGGCGGTACG                                                           |
| B-DOWN-2R     | CTCGGCGTGCCTTGAGGTGAGG                                                         |
| C-UP-2R       | CCCAGCAGCGAGTCTG                                                              |
| C-DOWN-2R     | CTGGCGGCTGAGGCGGAC                                                            |
| GROEL-2R      | CGGCGTCTGAGGCGGAC                                                             |
| N-3FLAG-A-P1  | CTACAGGGAGGACCCCATATGGAGTCAAGACCATGACCCAGG                                      |
| N-3FLAG-A-P2  | AAAGATCAGCATACATCGATTACAAGAGGATGACGAGGAAAG                                      |
| N-3FLAG-A-P3  | CAGCACCTTTGTCATCGTCTCTCTGCTGGAACCATGTAAGAGGCAAG                                  |
| N-3FLAG-A-P4  | TATAATCACCGTCTGAGTATTGCTGCTGAGGAC                                              |
| C-6HIS-A-F    | GCTGGCGGGGGGCGGACATCATCATCATCATCGAGGACGTT                                      |
| C-6HIS-A-R    | ATACCGGTACCTCAGATGATGATGATGATGAGGCCGCCGCCAGC                                   |
| N-6HIS-A-F    | GAGAGACCCCATATGACAGCGGGTTATATAAGATGACGAGC                                      |
| N-6HIS-A-R    | CATCGGATACGACATGAGCGGAGCCCAAAGGAAGGAC                                     |
| DR_0089-F     | CGCTGCGGTTTCCTCCC                                                              |
| DR_0089-R     | ACAAGCCGGCTGAGGCGGAC                                                          |
| DR_1998-F     | GGGCGGTGAGAAGCAGTG                                                              |
| DR_1998-R     | GTAGACGGGGGCTCTGCT                                                           |
| Name            | Sequence (5’→3’)                                      |
|-----------------|-------------------------------------------------------|
| DR_A0146-F      | GGCGCCGAGTACCAGCTGC                                   |
| DR_A0146-R      | CGTGCCAGGTTGAACCTAG                                  |
| DR_1279-F       | GGCAAGCTCGATGTCGTGCC                                  |
| DR_1279-R       | GGCGGCGTCTGGTTAGTGG                                  |
| DR_0644-F       | CGCATGATCGCCCAGGCG                                   |
| DR_0644-R       | GTGCAGATCAGCCAGCTG                                   |
| TRXI-BAMHI-F    | CAGCAATGGGGTCGCCGATCCATGAGTGAGTGACATCCTGACCTGTA      |
| TRXI-SALI-R     | TGCGGCGCAAGCTTGTCGACTCAAAGAGCTTGAGGTCGAGG           |
| TRXI-NCOI-F     | TAAGAACCGATATACCATGCGCATAGTGACATCGTCGACCTG          |
| TRXI-BAMHI-R    | ACGGAATCTCGAAATTCGGATCCCGAGGAAAGCTGGTTAGGTG         |
| SEP-NFLAG-F     | TAAGAAGGAGATATACCATGAGTGACTACAAAGGACCATGACG         |
| SEP-NFLAG-R     | ACGGAATCTCGAATTCGGATTCGAGCAGCAGCAGCAGC             |
| SEP-CFLAG-F     | TAGAGAACGATATACCAATGGTTAGGGTGACTCCGGTGACTGGACC     |
| SEP-CFLAG-R     | ACGGAGCTCAGATCCGCTGACCTGACTTCATCGTCGACTCCCTG       |
| DR_1343_F      | GGCTGGTTTCCGCATCCTC                                  |
| DR_1343_R      | GTTGACCGTCAAGGCTGCTTCC                               |
| MUT_TRX1-CFLAG-F | GCTTTGCATCCGAATTCCAGCTCCGCTCGAGCAAG               |
| MUT_TRX1-CFLAG-R | CGAATTCGGGATCGAAGCTGGTTAGGTGTTCCA                |
**Supplemental Table S2.** Proteases used for in-gel digestion of SEPs.

| Protease         | Ratio (W/W) | Digestion Buffer                                               | Digestion procedure                                                                                           |
|------------------|-------------|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| ArgC             | 1:100       | Incubation buffer: 50 mM Tris-HCl (pH 7.6–7.9), 5 mM CaCl_2, 2 mM EDTA | 1. Add the protease to incubation buffer, and then add the buffer to gel pieces. 2. Add activation buffer, 10×, to a final concentration of 1× 3. Incubate samples for 18 hours at 37 °C. |
| Chymotrypsin     | 1:50        | 100 mM Tris-HCl (pH 8.0), 10 mM CaCl_2 (pH 8.0)                 | 1. Add chymotrypsin to the gel piece with buffer. 2. Incubate samples for 18 hours at 25 °C. |
| LysC             | 1:50        | 50 mM Tris-HCl (pH 8.0)                                         | 1. Add LysC to the gel pieces with buffer. 2. Incubate samples at 37 °C for 18 hours. |
| LysN             | 1:50        | 50 mM Tris-HCl (pH 8.0)                                         | 1. Initiate digestion by adding 2.0 μg of LysN for a final enzyme to substrate ratio of 1:50. 2. Incubate samples at 37 °C for 18 hours. |
| Trypsin          | 1:50        | 50 mM NH_4HCO_3 (pH 7.8)                                        | Incubate samples at 37 °C for 18 hours. |
| Mirror-trypsin   | 1:50        | 20 mM Hepes (pH 7.5), 10 mM CaCl_2, 5 % ACN                      | 1. Add LysN to digestion buffer for a final enzyme ratio of 1:50 (W/W); add to gel pieces; and incubate for 4 hours at 37 °C. 2. Centrifuge the sample at 15,000 rpm for 10 min. 3. Add mirror-trypsin for a final enzyme to substrate ratio of 1:50 (W/W), and vortex at room temperature for 30-45 min. 4. Incubate for 18 hours at 37 °C. |
Supplemental Figure S1. A Schematic of different SEP discovery workflows used: PAGE+in-gel digestion+LC−MS/MS and MWCO+solution digestion+LC−MS/MS. The peptidome is separated by size using polyacrylamide gel electrophoresis (PAGE) or a 30 kDa MWCO filter (MWCO) and then analyzed directly by LC−MS analysis. B Venn diagram of identified SEPs used MWCO and PAGE.
Supplemental Figure S2. Predicted protein structure. Left panel, predicted protein structure of SEP068184 (Cyan), C-score=-1.15. Right panel, structural alignment of SEP068184 (Cyan) and human molybdopterin synthase complex (PDB: 5MPO) (Orange). TM-score=0.764, Cov=0.928.
Supplemental Figure S3. Protein–protein interaction networks of quantitative proteome. Nodes were significantly up- or down-regulated proteins in Panel A (adj.p-value<0.05, FC>2 or <0.5). Blue represents down-regulation, orange represents up-regulation.
Supplementary Data 1. The detailed information of the SEP candidates filtered at three stages.

Supplementary Data 2. Detailed information for quantitative peptidomics analysis of SEPs under irradiation and oxidative stress.

Supplementary Data 3. The detailed information of quantitative analysis of proteome and Co-IP/MS.

Supplementary Data 4. The detailed protein annotation information of whole proteome, quantitative proteome and Co-IP.

Supplementary Data 5. The annotated spectra of the SEPs containing one unique peptide.