A CRISPR-associated factor Csa3a regulates DNA damage repair in Crenarchaeon Sulfolobus islandicus

Liu, Zhenzhen; Sun, Mengmeng; Liu, Jilin; Liu, Tao; Ye, Qing; Li, Yingjun; Peng, Nan

Published in:
Nucleic Acids Research

DOI:
10.1093/nar/gkaa694

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC

Citation for published version (APA):
A CRISPR-associated factor Csa3a regulates DNA damage repair in Crenarchaeon Sulfolobus islandicus

Zhenzhen Liu1, Mengmeng Sun2, Jilin Liu1, Tao Liu1, Qing Ye1, Yingjun Li1 and Nan Peng1,*

1State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, 430070, P. R. China and 2Department of Biology, University of Copenhagen, DK-2200 Copenhagen N, Denmark

Received February 25, 2020; Revised August 06, 2020; Editorial Decision August 07, 2020; Accepted August 11, 2020

ABSTRACT
CRISPR—Cas system provides acquired immunity against invasive genetic elements in prokaryotes. In both bacteria and archaea, transcriptional factors play important roles in regulation of CRISPR adaptation and interference. In the model Crenarchaeon Sulfolobus islandicus, a CRISPR-associated factor Csa3a triggers CRISPR adaptation and activates CRISPR RNA transcription for the immunity. However, regulation of DNA repair systems for repairing the genomic DNA damages caused by the CRISPR self-immunity is less understood. Here, according to the transcriptome and reporter gene data, we found that deletion of the csa3a gene downregulated the DNA damage response (DDR) genes, including the ups and ced genes. Furthermore, in vitro analyses demonstrated that Csa3a specifically bound the DDR gene promoters. Microscopic analysis showed that deletion of csa3a significantly inhibited DNA damage-induced cell aggregation. Moreover, the flow cytometry study and survival rate analysis revealed that the csa3a deletion strain was more sensitive to the DNA-damaging reagent. Importantly, CRISPR self-targeting and DNA transfer experiments revealed that Csa3a was involved in regulating Ups and Ced-mediated repair of CRISPR-damaged host genomic DNA. These results explain the interplay between Csa3a functions in activating CRISPR adaptation and DNA repair systems, and expands our understanding of the lost link between CRISPR self-immunity and genome stability.

INTRODUCTION
CRISPR—Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated) systems are prokaryotic immune systems that protect Bacteria and Archaea against invasive viruses and plasmids (1,2). These adaptive immune systems are grouped into two major classes and six types (I−VI) (3). Class I CRISPR—Cas systems (Types I, III and IV) utilize the interference machinery composed of multiple Cas proteins, whereas class 2 systems (Types II, V and VI) employ a single Cas protein for interference. CRISPR immunity occurs in three functional stages, including adaptation (uptake of new spacers), crRNA biogenesis (transcription and processing of CRISPR RNA), and interference (nucleic acid-targeting and cleavage) (4,5). The conserved Cas1 and Cas2 proteins are essential in the adaptation stage to integrate spacers from invaders into CRISPR arrays (6). For instance, in the Sulfolobus islandicus subtype I-A system, the cas1 and cas2 genes, as well as the additional adaptation cas genes (cas3, cas5 and cas8) are transcriptionally activated for CRISPR adaptation via the CRISPR-associated factor Csa3a (7,8). However, CRISPR adaptation, in most cases, select spacers from host genomic DNA, in addition to the mobile genetic elements (MGEs) (8−10), thereby inducing self-immunity.

At the interference stage, CRISPR—Cas systems introduce DNA breaks at the target sites, both on the invasive and host genomic DNA (11). Damages on the genomic DNA caused by CRISPR interference, other environmental or endogenous factors could be repaired to maintain genome stability. There have been many studies on DNA damage repair in bacteria, archaea and eukaryotes, among which a unique regulatory network called DNA damage response (DDR) has been revealed (12,13). The regulatory mechanism of DDR governed by the RecA and LexA proteins is the best representative by the SOS response and is conserved in numerous bacteria (14). To ensure an effective and timely DNA damage repair in cells, the DDR mechanism shows some common characteristics, including rapid identification of DNA damage signals, after which a series of cellular events occur coordinately, for example, inhibition of DNA replication, cell cycle arrest and activating expression of DNA repair enzymes (15). The Crenar-
The archaeon *Sulfolobus* species encode specific DNA repair systems which transfer chromosome DNA between *Sulfolobus* cells for homologous recombination (16). Particularly, these systems rely on cell aggregation induced by the Ups (UV-inducible pili of *Sulfolobus*) post DNA damage (17,18), and the Ced (Crenarchaeal system for exchange of DNA) system directly function in the transfer of chromosomal DNA among cells (16). DNA transferred via the Ced system into cells is used as the donor for homologous recombination repair of the damaged DNA (16).

Ups- and Ced-mediated DNA repair in *Sulfolobus* species is efficiently regulated. Recent reports have revealed that Cdc6–2 is a central factor for activation of several DDR genes including *ups* and *ced* genes, DNA polymerase II (*dpo2*), homologous recombination repair genes, and the *tfb3* gene which encodes the secondary DDR regulator (19,20). Besides, Cdc6–2 represses expression of genes associated with cell division, DNA replication initiation, and genome segregation (21).

In *S. islandicus*, the CRISPR-associated factor Csa3a has been revealed to activate the expression of adaptation *cas* gene, CRISPR RNA and DNA repair genes, including helicase *herA*, nuclease *murA* and *dpo2* genes (8). However, it remains an open question on whether Csa3a regulates DDR pathways for DNA damage repair in *Sulfolobus*. Herein, we assessed whether Csa3a activates expression of *ups* and *ced* operons, as well as *cde6–2* and *tfb3* genes, and unveiled the crosstalk between the CRISPR–Cas system and DNA repair pathways in the Crenarchaeon *S. islandicus*.

**MATERIALS AND METHODS**

**Cell growth and DNA damage treatment**

*Sulfolobus islandicus* strains, including the wild-type (E233S) and the derived strains, are listed in Supplementary Table S1. These strains were cultured at 78°C in the SCV medium (basal medium supplemented with 0.2% sucrose, 0.2% casamino acids and 1% vitamin solution), or SCVU medium (SCV medium with addition of 20 μg/ml uracil at a final concentration) (22).

*Sulfolobus* genomic DNA damage treatment using 4-nitroquinoline-1-oxide (NQO), which mimics UV radiation, was performed as previously described (23). Briefly, NQO dissolved in DMSO solution was added to *Sulfolobus* cultures at the early exponential growth phase (OD<sub>600</sub> = 0.2) to a final concentration specified in each experiment. Cell concentration was determined at OD<sub>600</sub> during incubation, and cell samples were taken for aggregation assay, cell viability assay, RNA extraction, and flow cytometry analysis.

**Protein expression and purification**

Expression and purification of the Csa3a protein were conducted as described previously (7).

**Transcriptome analysis**

Strains (two biological repeats for each strain) for transcriptome analysis were cultured to log phase (OD<sub>600</sub> = 0.2). Thereafter, a 1 ml culture of each strain was transferred to 100 ml fresh SCVU medium in 250-ml flasks. Exponentially growing cultures of *S. islandicus* E233S (the wild-type strain, WT,) and Δ*csa3a* were diluted to OD<sub>600</sub> = 0.2 and cultured in the presence or absence of 2 μM NQO for 6 h. Then, total RNA was extracted using the Trizol reagent (Invitrogen, Carlsbad, CA, USA). Genomic DNA in the total RNA sample was removed using DNase I (Roche, Basel, Switzerland). The quality and quantity of purified total RNA were determined by measuring the absorbance at 260 and 280 nm using a NanoDrop ND-1000 spectrophotometer (Labtech, Wilmington, MA, USA). Total RNA integrity was verified by electrophoresis on a 1.5% agarose gel. A total of 3 μg RNA per sample was used as input material for cDNA library preparations. Sequencing libraries were generated using the NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer’s recommendations, and index codes were added to assign sequences to each sample. First-strand cDNA synthesis was performed using random hexamer primers and M-MuLV Reverse Transcriptase (RNase H<sup>-</sup>). Subsequently, second-strand cDNA synthesis was performed using DNA polymerase I and RNase H, followed by 15 cycles of PCR enrichment. Sequencing was performed with an Illumina HiSeq2000 instrument. Raw data were initially processed to obtain clean reads by eliminating adapter sequences and low-quality bases. Clean reads were aligned to the reference genome sequence of *S. islandicus* REY15A (GenBank Accession No. NC_017276). An index of the reference genome was built using Bowtie software v2.0.6, and paired-end clean reads were aligned to the reference genome using TopHat software v2.0.9. To count the number of reads mapped to each gene, HTSeq software v0.6.1 was used, following which the reads per kilobase per million mapped reads (RPKM) for each gene was calculated based on the length of the gene and the number of reads mapped to the gene. Each strain was sequenced in duplicate. To determine the expression level of each gene in different groups, transcript expression levels were expressed as the RPKM. Next, P-values were used to identify differentially expressed genes (DEGs) between the 2 groups using the chi-squared test (2 × 2), and the significance threshold of the P-value in multiple tests was set based on the false discovery rate (FDR). Furthermore, Fold-changes (log<sub>2</sub>[RPKM1/RPKM2]) were estimated according to normalized gene-expression levels. Threshold for DEGs was set at P-values < 0.01 and log<sub>2</sub> fold-change > 1 (FDR < 0.05). Transcriptome data were deposited in the SRA database under Accession PRJNA608153. The DEGs are listed in Supplementary Table S2.

**Electrophoretic mobility shift assay**

PCR was used to generate probes for electrophoretic mobility shift assay (EMSA) using one of the primers with 5′-end HEX-label (Supplementary Table S3). PCR products of the probes were first cloned into the T-vector as the template for inverse PCR to introduce mutations at the desired sites. Mutated probes were amplified from above plasmids introduced with mutations using one of the primers with 5′-end HEX-label. Then, the PCR products were purified in 6% native polyacrylamide gel electrophoresis (PAGE) for EMSA.
EMSA binding reactions (15 μl) containing 5 ng/μl of HEX-labeled probes and different concentrations of Csa3a protein (as described in the figure legends) were incubated for 20 min at 4°C in the binding buffer [20 mM Tris–HCl, pH 8.0, 50 mM KCl, 5% glycerol, 1 mM ethylenediaminetetraacetic acid, 1 mM dithiothreitol, 5 ng/μl poly(dI-dC)]. For the specific competition, increasing amounts of unlabeled specific probes were added to the reaction mixture. Thereafter, samples were loaded onto a 4% low melting agarose gel buffered with 1× TAE solution. DNA–protein complexes were separated at 200 V for 40 min and the resulting bands was detected using a FUJIFILM scanner (FLA-5100).

Localized surface plasmon resonance analysis

PCR was used to generate the probes for the localized surface plasmon resonance (LSPR) analysis using the primers listed in Supplementary Table S3. PCR products were purified through ethanol precipitation. The Csa3a protein was immobilized on the chip with different titer of probes in the mobile phase, as indicated in the figure legends. The kinetic parameters of the binding reactions were calculated and analyzed by Trace Drawer software (Ridgeview Instruments AB, Sweden) and One to One fitting model.

Flow cytometry analysis

The flow cytometry analysis was conducted as described previously (23). Briefly, 300 μl of Sulfolobus cells were fixed with 700 μl of absolute ethanol and stored at 4°C for 12 h. Then, fixed cells were collected through centrifugation at 2, 800 rpm for 20 min and washed with 1 ml of 10 mM Tris–NaCl buffer, pH 7.5, with 10 mM MgCl2. Cells were collected again and stained with 40 μg/ml ethidium bromide (Sigma-Aldrich, St. Louis, MO, USA) and 100 μg/ml mithramycin A (Apollo chemical, Tamworth, UK). We analyzed the stained cell samples in an Apogee A40 cytometer (Apogeeflow, Hertfordshire, UK) equipped with a 405 nm laser. A dataset of at least 60,000 cells were collected for each sample.

Cell aggregation assay

Cell aggregation in S. islandicus cultures was estimated through direct microscopy of cell aggregates in fresh cultures, as previously described (19). Briefly, each strain was cultured in the absence or presence of 2 μM NQO with a final concentration of 2 μM and incubated for 6 h. Then, using 1 ml of culture, cells were pelleted by centrifugation and re-suspended in 1 ml fresh SCVU medium. The resulting cell suspensions were serially diluted and plated using the two-layer plating method. Exactly 100 μl of the diluted samples were plated onto gelite plates in triplicate. Colonies appearing on plates after 7 days of incubation were counted, to determine CFUs/ml culture.

Plate titration experiment

The plate titration experiment of Sulfolobus was estimated by determining the density of the lawns on the plate. Exponentially growing cultures of S. islandicus (OD600 = 0.2) were treated with NQO to a final concentration of 2 μM and incubated for 6 h. Then, using 1 ml of culture, cells were serially diluted and 10 μl of the diluted sample was dripped on the SCVU plate and incubated at 78°C for 2 days.

DNA transfer and repair assays

DNA transfer assay was conducted as described previously (16), with modification. Targeting plasmid pTcas5 was constructed by cloning an oligonucleotide matching the protospacer on the cas5 gene in a mini-CRISPR cassette (Repeat-Spacer-Repeat) on the pSeSD plasmid. Then, 1 μg of pTcas5 plasmid was electroporated into 50 μl competent cells of S. islandicus wt (ΔpyrEFΔlacS) genes or the first genes of the DDR operons (SiRe1878: upsX, SiRe1879: upsE, SiRe1881: upsA, SiRe1316: cedA1, SiRe1857: cedB, SiRe1717: tfb3 and SiRe1231: cdc6–2) that were up-regulated in the csa3a overexpression cells. Promoter fragments of these genes were amplified with S. islandicus REY15A genomic DNA using Phanta DNA Polymerase (Vazyme, Nanjing, China), and the primers listed in Supplementary Table S3. The PCR products (ca. 200 bp) were purified using the Cycle-Pure kit (Omega Bio-Tek, USA). Purified DNAs were digested and inserted into the pSe-lacS vector (25) to yield the reporter plasmids: pSe-upsX-lacS, pSe-upsE-lacS, pSe-upsA-lacS, pSe-cedA1-lacS, pSe-cedB-lacS, pSe-tfb3-lacS and pSe-cdc6–2-lacS (Supplementary Table S1).

These reporter gene plasmids were electroporated into S. islandicus wild-type (E233S) and Δcsa3a strains, respectively. Three single colonies of each transformant were selected and cultured for 6 h in SCV either in the presence or absence of 2 μM NQO. Cell mass was collected for each strain and used to prepare the cellular extracts. The protein content of the cellular extracts was determined by microBCA Protein Assay Reagent (Thermo Scientific), whereas the β-galactosidase activity was determined as previously described (24).

Cell viability analysis

We estimated the cell viability of Sulfolobus cultures by determining their colony formation units (CFU)/ml culture. Exponentially growing cultures of S. islandicus (OD600 = 0.2) were treated with NQO to a final concentration of 2 μM and incubated for 6 h. Then, using 1 ml of culture, cells were pelleted by centrifugation and re-suspended in 1 ml fresh SCVU medium. The resulting cell suspensions were serially diluted and plated using the two-layer plating method. Exactly 100 μl of the diluted samples were plated onto gelite plates in triplicate. Colonies appearing on plates after 7 days of incubation were counted, to determine CFUs/ml culture.

Report gene assay

Reporter plasmids were constructed using the Sulfolobus–E. coli shuttle vector pSeSD with the S. solfataricus galactosidase gene (lacS) as the reporter gene (24). For this experiment, we selected the promoters of the single DDR operons or the first genes of the DDR operons (SiRe1878: upsX, SiRe1879: upsE, SiRe1881: upsA, SiRe1316: cedA1, SiRe1857: cedB, SiRe1717: tfb3 and SiRe1231: cdc6–2) that were up-regulated in the csa3a overexpression cells. Promoter fragments of these genes were amplified with S. islandicus REY15A genomic DNA using Phanta DNA Polymerase (Vazyme, Nanjing, China), and the primers listed in Supplementary Table S3. The PCR products (ca. 200 bp) were purified using the Cycle-Pure kit (Omega Bio-Tek, USA). Purified DNAs were digested and inserted into the pSe-lacS vector (25) to yield the reporter plasmids: pSe-upsX-lacS, pSe-upsE-lacS, pSe-upsA-lacS, pSe-cedA1-lacS, pSe-cedB-lacS, pSe-tfb3-lacS and pSe-cdc6–2-lacS (Supplementary Table S1).

These reporter gene plasmids were electroporated into S. islandicus wild-type (E233S) and Δcsa3a strains, respectively. Three single colonies of each transformant were selected and cultured for 6 h in SCV either in the presence or absence of 2 μM NQO. Cell mass was collected for each strain and used to prepare the cellular extracts. The protein content of the cellular extracts was determined by microBCA Protein Assay Reagent (Thermo Scientific), whereas the β-galactosidase activity was determined as previously described (24).

Plate titration experiment

The plate titration experiment of Sulfolobus was estimated by determining the density of the lawns on the plate. Exponentially growing cultures of S. islandicus (OD600 = 0.2) were treated with NQO to a final concentration of 2 μM and incubated for 48 h. Then, using 1 ml of culture, cells were serially diluted and 10 μl of the diluted sample was dripped on the SCVU plate and incubated at 78°C for 2 days.
or Δcsa3a (ΔpyrEFΔlucSΔcsa3a) cells. After electroporation, 1 ml preheated medium was added into the culture containing the transformed cells. The mixed cultures were incubated with or without 2 ml of preheated Δcas5 (ΔpyrEFΔlucSΔcas5) mating partner cells (OD600 = 0.5) for 2 h at 78°C, then, plated on the SCV medium without uracil at 78°C for 6 days. We calculated increased folds of transformation efficiencies of wt or Δcsa3a cells incubated with Δcas5 partner cells compared with that of wt or Δcsa3a cells (Ewt::pTcas5×Δcas5/Ewt::pTcas5 or EΔcsa3a::pTcas5×Δcas5/EΔcsa3a::pTcas5; E: transformation efficiency), respectively. Deletion of the cas5 gene locus on the chromosomes of the single colonies of wt::pTcas5 × Δcas5 or Δcsa3a::pTcas5 × Δcas5 conjugates were determined by PCR analysis.

RESULTS

Csa3a acted as a transcriptional activator for the DDR genes

The S. islandicus strain REY15A carries a subtype I-A adaptation strain consisting of the cas1 (SiRe_0760), cas1 (SiRe_0761), cas2 (SiRe_0762) and cas4 (SiRe_0763) genes, which are regulated by a factor encoded by the CRISPR-associated csa3a gene (Figure 1A) (7). Besides, this strain carries an ups operon and a ced cluster (Figure 1B and C), encoding pili which aid in cell aggregation (18) and encoding a membrane system for DNA transfer, thereby mediating DNA damage repair via homologous recombination (16).

Upon re-analysis of our previous transcriptome data (8), we found that most of the DDR genes, including the ups and ced genes, as well as the cdc6-2 and tfb3 genes, were significantly up-regulated in the csa3a overexpression strain (Supplementary Table S4). In this study, we revealed that most of these genes were correspondingly down-regulated in the csa3a deletion strain based on the transcriptome data (Table 1). These results inferred that the Csa3a factor acts as a transcriptional activator for the DDR genes in S. islandicus.

A previous report indicated that DDR genes were induced by NQO reagent (23), we observed a similar phenomenon based on our transcriptome data (Table 1). NQO-induced expression of DDR genes is much stronger than Csa3a induced (Table 1). NQO also induced DDR expression in the csa3a gene deletion strain (Table 1), indicating Csa3a might not be essential for DDR induction upon NQO treatment. However, transcriptome data for the csa3a deletion strain and the wild-type strain both treated with NQO revealed that transcription of ups4 and tfb3 genes (Table 1) were significantly down-regulated in csa3a deletion strain (Table 1). The later was essential for the regulation of DDR genes (19). Conclusively, these findings indicated Csa3a might play an important, but not essential, role in the regulation of DDR genes in S. islandicus.

Csa3a specifically bound to the DDR gene promoters

To assess whether Csa3a directly bound to the promoters of the DDR genes to regulate their expression, DNA fragments of the upsX (SiRe_1878), upsE (SiRe_1879), upsA (SiRe_1881), cedA1 (SiRe_1316), cedB (SiRe_1857), cdc6-2 (SiRe_1231) and tfb3 (SiRe_1717) promoters were used as probes for EMSA and LSPR experiments. In our previous investigation, we found that Csa3a specifically bound to cas1 promoter and the leader sequence, and deletion of or mutations at the binding sites completely abolished the binding (7,8). Here, the truncated leader sequence (−94 to −15, relative to the first repeat) with deletion of the confirmed Csa3a binding site (8) was used as the non-specific probes, and no shift was found in EMSA experiments using Csa3a protein and the non-specific probe (Supplementary Figure S1). Then, we assessed the Csa3a-binding ability using the full-length promoters of the upsA (198 bp), upsX (227 bp), upsE (300 bp), cedA1 (214 bp), cedB (198 bp), cdc6-2 (200 bp) and tfb3 (200 bp), upstream of the start codon since the transcriptional start sites were not determined yet. An increase in signal intensity of the retarded bands was parallel with increasing Csa3a amounts (40, 80 or 120 ng/μl) in each EMSA experiment (Figure 2A). On the other hand, the signal of retarded band for each EMSA experiment was completely abolished for all tested promoters in the presence of 2- or 4-fold excess of unlabelled specific competitor DNA (cold probe) (Figure 2A).

Notably, a DNA motif similar to the UV-responsive motif (26) was identified for all tested promoters. The mutated promoters with transversion mutation at the UV-responsive motif and its flanking region formed a weak shift band with the Csa3a regulator (40, 80 or 120 ng/μl) in EMSA experiment (Figure 2A). Further, the LSPR results showed different Kd values for the interaction between Csa3a and each wild-type promoter or the mutant promoters (with mutations at the UV-responsive motif and the flanking sequence) (Figure 2B). All LSPR analysis revealed that Csa3a strongly bound to the full-length promoters with Kd values ranging from 0.29 to 1.30 μM (Figure 2B). Generally, the Kd values for the binding between the mutant promoters and Csa3a showed a lower affinity (1–2 orders of magnitude) compared to that between the wild-type promoters and Csa3a (Figure 2B). All LSPR raw data were summarized in Supplementary Table S5. It should be noticed that we previously demonstrated that Csa3a strongly bound a palindromic DNA sequence on the cas1 promoter and the leader sequence (7,8). However, this DNA motif was not identified on the DDR gene promoters. This might explain the relative weak interaction between the Csa3a factor and the DDR gene promoters, and explain that mutations have a slight effect on their binding in the EMSA and LSPR analyses. Taken together, the EMSA and LSPR results and the transcriptome data indicated that Csa3a specifically binds to these DDR gene promoters to activate their transcription.

Deletion of the csa3a gene significantly reduced the promoter activities of the DDR genes

To assess whether Csa3a could regulate the DDR gene expression in vivo, promoter fragments of these genes (upsX, upsE, upsA, cedA1, tfb3 and cdc6-2) were used in the reporter gene assay (Figure 3A). These promoters were cloned at the immediate upstream of the lacS gene, encoding a β-galactosidase, to control its transcription in an
Deletion of the csa3a gene reduced DNA damage-induced cell aggregation

To assess the impact of the Csa3a regulator on Ups-mediated cell aggregation, *S. islandicus* wild-type (WT) and the *csa3a* deletion cells were cultured in presence or absence of 2 μM NQO for 12 h, after which cell samples were examined microscopically. Cell aggregation was observed for neither wild-type strain nor Δcsa3a strain at all tested time points without NQO treatment (Figure 3B). However, Δcsa3a gene deletion did not significantly impact the expression of these genes after NQO treatment, except for the *cd6-2* gene (Figure 3B). All raw data for reporter gene assay are summarized in Supplementary Table S6. These findings inferred Csa3a is a key factor in the regulation of the DDR gene expression probably without or with less environmental stresses (e.g. UV radiation or NQO treatment). Moreover, it was suggested that Csa3a could directly regulate the DDR genes or indirectly regulate them via regulation of the *cd6-2* gene, which was the central factor that regulated DDR gene expression (21).

Deletion of the *csa3a* gene reduced cell viability upon NQO treatment

A recent study revealed that Cdc6-2 affect cell viability in the Ups- and Ced-mediated DDR pathway (21). In this study, we cultured *S. islandicus* wild-type, Δcsa3a, and Δcd6-2 strains for 48 h in the medium with or without 2 μM NQO. Then, cells were analyzed by flow cytometry. DNA-less cells were observed for almost all the strains in the NQO treated cultures, and the population of DNA-less cells was proportional to the time of NQO exposure (Figure 5A). Significantly, the Δcd6-2 strain produced many cells with one chromosome (G1 cells), while the population of DNA-less cells with the Δcsa3a strain was more than that in the wild-type and Δcd6-2 cells from 6 to 12 h post NQO treatment (Figure 5A). Moreover, in the *csa3a* deletion and wild-type cells, there were a few G1 cells and more G2 cells.

Table 1. Transcriptional changes (log2-fold) of DNA damage response genes in *S. islandicus* according the transcriptome data

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Annotation</th>
<th>Δcsa3a vs wt</th>
<th>Δcsa3a+NQO vs WT</th>
<th>WT+NQO vs wt</th>
<th>Δcsa3a+NQO vs wt+NQO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA transfer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiRe1,1878</td>
<td>upsX</td>
<td>-</td>
<td>6.29</td>
<td>5.54</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1879</td>
<td>upsE</td>
<td>-2.07</td>
<td>7.81</td>
<td>6.35</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1880</td>
<td>upsF</td>
<td>-1.30</td>
<td>6.71</td>
<td>5.65</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1881</td>
<td>upsA</td>
<td>-2.21</td>
<td>7.54</td>
<td>6.54</td>
<td>-1.16</td>
</tr>
<tr>
<td>SiRe1,1882</td>
<td>upsB</td>
<td>-1.79</td>
<td>7.39</td>
<td>6.49</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1316</td>
<td>cedA1</td>
<td>-</td>
<td>6.59</td>
<td>6.38</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1317</td>
<td>cedA2</td>
<td>-0.96</td>
<td>6.64</td>
<td>6.06</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1857</td>
<td>cedB</td>
<td>-0.93</td>
<td>5.75</td>
<td>5.47</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transcriptional regulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiRe1,1231</td>
<td>cdc6-2</td>
<td>-</td>
<td>3.46</td>
<td>3.53</td>
<td>-</td>
</tr>
<tr>
<td>SiRe1,1717</td>
<td>tjb3</td>
<td>-1.84</td>
<td>6.55</td>
<td>5.89</td>
<td>-1.12</td>
</tr>
<tr>
<td>SiRe0,0764</td>
<td>csa3a</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Note:* Unchanged; NA: not apply.

Significance: log2-fold change > 1.
Figure 2. Analysis of Csa3a binding with the promoters of ups, ced, tfb3 and cdc6–2 genes. (A) EMSA analysis of Csa3a binding to the upsX (SiRe_1878), upsE (SiRe_1879), upsA (SiRe_1881), cedA1 (SiRe_1316), cedB (SiRe_1857), tfb3 (SiRe_1717) and cdc6–2 (SiRe_1231) promoters. For binding assays using the wild-type promoters or the mutated promoters, each reaction contained the 5′-end HEX labeled probes: 5 ng/μL, poly(dI-dC): 5 ng/μL, and Csa3a protein: 40, 80 or 120 ng/μL. For the specific competition assay, each reaction contained the 5′-end HEX labeled probes: 5 ng/μL, poly(dI-dC): 5 ng/μL, and Csa3a protein: 120 ng/μL, and unlabeled specific competitor (cold probe): 10 or 20 ng/μL. The probe location and mutated region on each promoter are indicated in relation to the ATG codon of each open reading frame. The wild-type and mutated sequences are indicated and the UV-responsive element (26) is boxed. P and PM: the wild-type probes and the mutated probes used in EMSA and LSPR experiments, respectively. W: precipitation at loading wells; S: shift; F: free probe. (B) LSPR analysis of fixed Csa3a protein on the chip to bind the promoters and their mutants used in (A). The concentrations of probes used for the analysis are shown. $K_D = \text{mean±standard derivations of three independent experiments.}$
with two chromosomes post NQO treatment (Figure 5A), indicating these strains exhibited stronger DNA repair activity. However, nearly all cdc6-2 deletion cells transformed into the DNA-less cells, showing cell death under the NQO treatment (Figure 5A). Importantly, more DNA fragments were observed in csa3a deletion strain from 6 to 48 hours post NQO treatment compared with the wildtype (Figure 5A), indicating a lower DNA repair efficiency in the csa3a deletion mutant.

The growth curves of the strains treated the same in Figure 5A showed that the culture containing 2 μM NQO could completely inhibit the growth of the Δcsa3a and Δcdc6-2 strains, whereas, for the wild-type strain, no obvious growth defect was observed (Figure 5B, Supplementary Table S8). From these findings, the deletions of csa3a and cdc6-2 genes showed hypersensitivity to the NQO treatment. The plate titration experiment also confirmed that the amount of viable cells in the Δcsa3a culture treated with NQO for 24 h with 10-fold dilutions was lower than that of wild-type strain with the same treatment and at the same dilutions (Figure 5C). We also evaluated the sensitivity of the wild-type and Δcsa3a strains to NQO by determining their survival rate on plates after 2 μM NQO treatment for 6 h. The number of viable cells, estimated by their colony formation units (CFU), revealed that NQO-treated cultures exhibited a strong reverse correlation to the drug concentration for both wild-type and Δcsa3a strains (raw data shown in Supplementary Table S9). However, the viable rate of the Δcsa3a strain was significantly lower than that of the wild-type strain (Figure 5D). In summary, the csa3a gene deletion significantly elevated the sensitivity of S. islandicus cells to the DNA damage reagent.

Csa3a is important for regulation of Ups- and Ced-mediated repair of CRISPR-damaged host genomic DNA

To study whether the Csa3a factor is involved in the regulation of Ups- and Ced-mediated repair of CRISPR-damaged host genomic DNA, we designed a targeting plasmid (pTcas5) against cas5 gene in the wt (ΔpyrEF ΔlacS) or Δcsa3a (ΔpyrEF ΔlacSΔcsa3a) cells, then, used a cas5 deletion strain (ΔpyrEF ΔlacSΔcas5) as the mating partner cells to provide repair donor DNA. Plasmid pTcas5 was electroporated into wt or Δcsa3a cells (resulting in wt::pTcas5 or Δcsa3a::pTcas5), after which these cells were incubated either in presence or absence of Δcas5 cells (Figure 6A). The genomic DNA of transformants carrying pTcas5 plasmids was targeted at the cas5 gene locus (Figure 6B). However, if Ups and Ced systems import genomic DNA of Δcas5 mating partner cells into the wt or Δcsa3a cells carrying pTcas5 plasmid (the conjugants of wt::pTcas5 × Δcas5 or Δcsa3a::pTcas5 × Δcas5), the damaged DNA was repaired through homologous recombination pathway (Figure 6B). Moreover, since the donor DNA carried a deletion at the

**Figure 3.** Analysis of promoter activities using the reporter gene system in the strains with or without csa3a gene. (A) The reporter plasmid used in this study. Promoters ofDDR genes were cloned immediately upstream of lacS gene which encodes β-galactosidase. (B) The specific β-galactosidase activities for the tested promoters in wild-type (S. islandicus E233S) or csa3a deletion strains with or without treatment with 2 μM NQO for 6 h. Error bars: standard derivations of three independent experiments. Statistical significance: *P < 0.05, ***P < 0.001, two-way ANOVA and Dunnett.
Figure 4. Deletion of csa3a gene reduced DNA damage-induced cell aggregation. (A) Microscopic analysis of cell aggregates in samples taken from the cultures of the wild-type (E233S) and Δcsa3a mutants at different time points with or without NQO treatment. Red arrows indicate the example of cell aggregates, and the aggregate with more than 30 cells in the E233S sample at 12 h post NQO treatment was circled. (B) Quantification data of cell aggregates in the cell samples shown in panel A. At least 1000 cells were analyzed for each sample. Error bars: standard derivations of three independent experiments. Statistical significance: *P < 0.05, ***P < 0.001, two-way ANOVA and Dunnett.
Figure 5. The csa3a deleted strain showed higher sensitivity to NQO. (A) Cell cycle profiles of wild-type, Δcsa3a and Δcdc6-2 cultures. DNA contents were divided into 256 arbitrary points on the X-axis, and cell counts (Y-axis) were obtained for each point and plotted against the DNA content. Each sample was grown in SCVU in the presence (denoted as +NQO), or absence (–NQO) of 2 μM NQO for 6 h. DNA-less cells (L), cells containing one chromosome (1), and cells containing two chromosomes (2). Red arrows indicate the DNA-less cells. (B) Growth curves based on absorbance at 600 nm. (C) Plate titration of cells. Each strain was grown in the absence or presence of NQO for 48 h, and a series of dilutions were prepared for each sample which was plated for 12 h. WT: S. islandicus E233S; Δcsa3a: csa3a deletion strain; Δcdc6-2: cdc6-2 deletion strain. Error bars: standard derivations of three independent experiments. (D) Survival rates of S. islandicus wild-type (E233S) and Δcsa3a strains after NQO treatment. Exponentially growing strains were treated with 2 μM NQO for 6 h. Cell samples were plated on NQO-free SCVYU plates for determination of colony formation units (CFU)/mL culture. Survival rate = (CFU/mL of NQO treated cells)/(CFU/mL of non-treated cells). Statistical significance: ***P < 0.001, two-way ANOVA and Dunnett.

csa5 gene locus, the repaired DNA would not be further targeted (Figure 6B). Transformation of the self-targeting plasmid pTcas3 resulted in 10²- to 10³-folds lower transformation efficiency compared with the transformation of the control plasmid pSeSD (Supplementary Table S10). Transformation efficiency of wt::pTcas5 resulted in 10²- to 10³-folds lower transformation efficiency compared with the transformation of the control plasmid pSeSD (Supplementary Table S10). While, transformation efficiency of wt::pTcas5 × Δcsa3a: pTcas5 were 3.05–4.40 × 10² and 3.30 × 10² cfu/μg plasmid DNA, respectively (Supplementary Table S10). Therefore, for both wt and Δcsa3a cells transformed with pTcas5, their transformation efficiencies increased by more than 10-folds when incubated with Δcas5 mating partner cells (Figure 6C). Together with previous reports that deletion of ups and ced genes failed to transfer DNA for DNA damage repair in Sulfolobus cells treated with UV light (16,27), our result indicated that the Ups and Ced systems mediated DNA damage repair at the targeted csa5 gene locus. Notably, csa3a gene deletion showed significantly lower repair efficiency (Figure 6C), an indication that the csa3a gene is critical for regulation of Ups- and Ced-mediated repair of CRISPR-damaged host genomic DNA.

To reveal on DNA exchange and homologous recombination, PCR was performed on the colonies of conjugants obtained in mixtures of wt::pTcas5 × Δcsa5 or
Deletion of csa3a gene reduced the efficiency of Ups- and Ced-mediated repair of CRISPR-damaged host genomic DNA. (A) schematic diagram of Ups- and Ced-mediated DNA repair analysis. Plasmid pTcas5 carrying a mini-CRISPR (Repeat-Spacer-Repeat) with the spacer against cas5 gene was electroporated into S. islandicus wt (ΔpyrEFΔlacS) and Δcsa3a (ΔpyrEFΔlacScsa3a) cells (step 1). After electroporation, the transformed cells were incubated with or without Δcas5 (ΔpyrEFΔlacSΔcas5) mating partner cells for 2 h at 78°C (step 2) and then plated in the plates of SCV medium without uracil (step 3). (B) schematic diagram of homologous recombination of CRISPR-damaged genomic DNA in wt or Δcsa3a cells carrying the pTcas5 self-targeting plasmid and containing the donor DNA imported from Δcas5 cells. Subtype I-A Cascade complex targeted site at cas5 gene locus is shown in red. The cas5 gene indicates as a blue arrow. Deletion of the cas5 gene is indicated as a dash line. Double-crossover between the homologous regions upstream and downstream of cas5 gene is shown. F and R represent the primers used for PCR amplification of the cas5 gene locus on the chromosome of the colonies. (C) fold increases of the transformation efficiencies of wt or Δcsa3a cells incubated with Δcas5 mating partner cells compared with that of wt or Δcsa3a cells (Ewt[wt:pTcas5×Δcas5]/Ewt[wt:pTcas5] or EΔcsa3a[wt:pTcas5×Δcas5]/EΔcsa3a[wt:pTcas5]; E: transformation efficiency) were calculated, respectively. Statistical significance: *P < 0.05, two-way ANOVA and Dunnett. (D) PCR amplification of the cas5 gene locus on the chromosomes of 20 randomly selected wt::pTcas5 × Δcas5, Δcsa3a::pTcas5 × Δcas5, wt::pTcas5 and Δcsa3a::pTcas5 single colonies, respectively, on SCV plates without addition of uracil. L, DNA ladder; a, PCR control using S. islandicus REY15A genomic DNA as the template; b, PCR control using the Δcas5 strain genomic DNA as the template; 1–20, PCR amplification of cas5 gene locus using 20 randomly selected colonies, respectively.

Discussion

Environmental factors, such as UV and other types of radiation, can cause DNA damage in the form of strand breaks, which are particularly lethal and can be mutagenic. The hyperthermophilic archaean Sulfolobales genus encodes the Ups and Ced, the special DNA damage repair system (16). The ups operon (cedA1, ceda, ceda2 and cedB) encodes transmembrane proteins (Figure 1C). For instance, the ceda1 and ceda2 encode two small transmembrane proteins, ceda encodes a larger transmem-
A proposal for Csa3a-centered network of DNA damage response for CRISPR immunity in *S. islandicus*. MGE invasion activates the expression of the global regulator Csa3a (32). Csa3a in turn activates expression of DDR genes, including *ups* and *ced* operons, and triggers CRISPR adaptation and immunity against invaders (7, 8). However, a subset of host DNA derived spacers would guide self-immunity against the genomic DNA. The activated Ups system aggregates cells and the Ced system transfers DNA from adhered cells into the target cells. The transferred DNA was used as the donor DNA for homologous recombination (HR) to remove newly integrated spacers and to repair the DNA breaks, resulting in intact genomic DNA. Ups: UV-responsive pili of *Sulfolobus*; Ced: Crenarchaeal system for exchange of DNA. Diamond: the CRISPR repeat sequence.

Brane protein, whereas *cedB* encodes a HerA/VirB4 homolog (16). UV-induced DNA damage up-regulates these *ups* genes, resulting in pili formation thereby causing cells to aggregate (17, 18, 27–30). Notably, the aggregations are species-specific mating pairs (17). VirB4/HerA homolog of the Ced system is, in most cases, highly up-regulated in *S. islandicus* after DNA damage (19, 21), and is induced by overexpression of the CRISPR-associated factor Csa3a (8) (Supplementary Table S4). Furthermore, VirB4-ATPases are associated with conjugative type-IV secretion systems in bacteria, where they are essential for DNA transfer (16). Homologs of these genes are specific for the Crenarchaeota and can be found in most species of Sulfolobales, Acidilobales, and Desulfurococcales (16). Otherwise, this system functions as a DNA importer and plays an important role in DNA homologous recombination repair (16). CedA proteins are thought to assemble a membrane pore through which the DNA can be transferred (16). Membrane-bound ATPase CedB presumably binds the DNA and energizes the translocation (16). Once the DNA is introduced, it can potentially repair the extensive DNA damage caused by UV through homologous recombination (16, 17, 31).

Ups-Ced-mediated DNA damage repair is regulated during exposure to UV radiation. Since DNA transfer requires both Ups and Ced proteins, it makes more sense to use a common factor to regulate the two systems. In a previous study, *S. islandicus* Cdc6–2 protein was confirmed to specifically bind to the promoters of the *ups* and *ced* genes, thereby facilitating the expression of these genes (21). However, a higher Cdc6–2 expression level is essential but not sufficient to trigger DDR regulation in this archaean (21). TFB3, a truncated version of the archaeal transcription factor B family proteins, has been found to regulate both Ups and Ced systems (19, 20), downstream of Cdc6–2 regulation in the DDR network in *S. islandicus* (21).

In our previous work, we identified Csa3a as a transcriptional regulator that activates the expression of adaptation *cas* genes (7) and enhances transcription of CRISPR RNAs (8). Besides, the expression of the Csa3a activator is induced by an unknown mechanism upon invasion by mobile genetic elements (32). Therefore, Csa3a was thought to be a dedicated factor specific to CRISPR—Cas regulation (33). However, Csa3a-triggered CRISPR adaptation integrates a few (7.0%) host spacers into CRISPR arrays (8), inducing self-immunity. In addition, the adaptation of host self-DNA occurred in other model systems. For example, in *de novo* spacer acquisition process, 32%, 16% and 22.8% or 1.8% (for induction or non-induction of Cas1, Cas2 expression, respectively) of new spacers have been derived from host genomic DNA in *S. thermophilus* subtype II-A (10), *Pectobacterium atrosepticum* subtype I-F (34) and *E. coli* subtype I-E (9) systems. Even in primed acquisition processes, 0.01–0.03% of new spacers have been derived from host genomic DNA in *E. coli* subtype I-E (35) and *P. atrosepticum* subtype I-F (34) systems. This raises an important question on how the cells evade self-immunity guided by self-derived spacers. Of note, CRISPR—Cas systems have developed diverse mechanisms to evade autoimmunity. For instance, bacterial and archaean genomes encode anti-CRISPR proteins to inhibit CRISPR immunity (36–39), therefore, reduce the threats to host genomic DNA.

In this study, our results revealed that the CRISPR-associated factor Csa3a of *S. islandicus* could specifically bind to the promoters of DDR genes, including *ups*, *ced*,...
tfb3 and cdc6−2 in vitro (Figure 2), and deletion of the csa3a gene significantly suppressed the expression of these genes (Table 1 and Figure 3), reducing the formation of cell aggregates (Figure 4). As a result, deletion of csa3a gene increased the sensitivity of cells to the DNA-damaging reagent NQO (Figure 5). Analysis of the transcriptome data revealed that DDR genes were lower in csa3a deletion strain compared with the wild-type without NQO treatment (Table 1), indicating Csa3a activates DDR expression. Notably, all DDR genes were significant higher in Δcsa3a cells treated with NQO than in cells not treated with NQO (Table 1). This implies that the Cdc6-2 and TFB3-signaling pathway became dominant when cells were treated with a DNA-damaging reagent (19,21). Moreover, Csa3a regulated the expression of cdc6-2 and tfb3 genes (Table 1 and Figure 2), suggesting Cdc6-2 and TFB3 are involved in Csa3a-mediated regulation of DDR genes.

Our further experiments confirmed that CRISPR-damaged host genomic DNA was repaired through mating with the Sulfolobus cells lacking the CRISPR targeting region (Figure 6), inferring the involvement of the Ups and Ced systems in this process. Importantly, deletion of csa3a gene resulted in reduced repair efficiency (Figure 6), indicating the Csa3a factor played an important regulatory role in the DNA damage repair process. Moreover, loss of newly adapted spacer during virus infection was detected in S. islandicus (40), implying the single strand nicks generated at leader-repeat region during spacer acquisition could also induce homologous recombination probably through the Ups-Ced pathway. Therefore, we propose a model for this regulation (Figure 7): (i) MGE invasion activates Csa3a expression (32), (ii) Csa3a triggers CRISPR adaptation and CRISPR RNA transcription (7,8), (iii) self-sparers guide immunity against host DNA, (iv) Csa3a activates the expression of Ups, Ced and DNA repair genes, (v) Ups system mediated cell aggregations and Ced system transfer genomic DNA into damaged cells (16–18,28,30), (vi) homologous recombination repair occurred at the DNA breaks of the target sites and leader-proximal regions, resulting in repair of targeted sites and loss of new spacers. This model well explains the interplay between the Csa3a functions in triggering CRISPR adaptation and activation of the DNA repair systems, and expands our understanding of the lost link between CRISPR self-immunity and genome stability.

DATA AVAILABILITY
Transcriptome data were deposited in the SRA database under Accession PRJNA608153.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank all co-authors’ kindly cooperation, especially in the special period of coronavirus outbreak in Wuhan city.

FUNDING
National Natural Science Foundation of China [91751104, 31671291 to N.P., 31900400 to T.L.]; National Postdoctoral Program for Innovative Talents [BX20180112 to T.L.]; Fundamental Research Funds for the Central Universities [2662019PY028 to N.P.]. Funding for open access charge: the National Natural Science Foundation of China [31671291].

Conflict of interest statement. None declared.

REFERENCES


