

参赛学生姓名: Li Sheng, Guanhong Liu

中学: Beijing 101 middle school

省份: Beijing

国家/地区: China/Beijing

指导老师姓名: Sun Wei

指导老师单位: Institute of Biophysics,

Chinese Academy of Sciences

论文题目: CRISPR-Cas-Based Detection of

Bacterial Nucleic Acids in Environmental

Water Samples

CRISPR-Cas-Based Detection of Bacterial Nucleic Acids in Environmental Water Samples

Li Sheng, Guan hong Liu

Abstract

Bacteria are widely distributed in natural environments, and their abundance and composition are closely related to local environmental conditions. Detecting bacterial DNA in water can help us to understand the microorganisms present in different water environments.

Over nearly one year, we learned and practiced molecular biology and biochemistry techniques, including molecular cloning, protein expression and purification and PCR et al. We then used these techniques to develop a CRISPR-Cas12a method for detecting bacterial DNA in environmental water samples. Water samples were collected from multiple sites at Beijing 101 Middle School, *Yuanmingyuan*, and the Olympic Forest Park. Some sites were sampled more than once to test the method with water collected at different times and locations. Because the water contained sediment and suspended material, we tested different ways to concentrate microorganisms before DNA extraction. Sequential filtration followed by collection on a 0.22- μm membrane was used to concentrate microorganisms from the water. The extracted DNA was then amplified by PCR.

For CRISPR-based DNA detection, we prepared AsCas12a protein and crRNAs targeting selected 16S rRNA gene markers. We tested a general gut-associated bacterial marker, a human gut-associated marker, and two different avian gut-associated markers. The Cas12a-crRNA complex was used to recognize the target DNA containing the PAM sequence within the 16S rRNA gene, and an ssDNA fluorescent reporter was used to generate the detection signal.

We tested water samples collected from multiple sites at three locations and found different patterns of bacterial DNA markers among the samples. These patterns were generally consistent with differences in the surrounding environments and human and animal activities at the sampling sites. To improve the reliability of the results, we carried out repeated sampling and adjusted the sampling sites and collection methods based on the first round of experiments. Our results suggest that AsCas12a-based detection can be used to detect selected bacterial DNA markers in urban environmental waters.

Our study shows that the molecular biology methods we learned can be combined into a workflow for detecting selected bacterial DNA in environmental water. The results also show that sample processing and target selection can strongly affect the detection results. More samples and bacterial markers will be needed to better understand the microbial communities

in these water environments.

Key words: CRISPR-Cas12a, Bacterial DNA Detection, Microbial source tracking, PCR, Environmental water samples.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Table of contents

CRISPR-Cas-Based Detection of Bacterial Nucleic Acids in Environmental Water Samples .	1
Abstract.....	1
1. Introduction.....	5
1.1 Detecting Bacteria in Water Is Essential	5
1.3 DNA detection by CRISPR-Cas12a	5
1.4 Nucleic acid extraction and processing.....	6
1.5 What We Did in This Study	7
2. Methodology	7
2.1 Cloning, expression and purification of Cas12a proteins.....	7
2.1.1 Cloning of Cas12a genes into pET28a vectors.....	7
2.1.2 Overexpression of Cas12a proteins	8
2.1.3 Purification of Cas12a proteins.....	8
2.2 crRNA preparation and DNA synthesis.....	9
2.3 Water sampling	11
2.4 Preliminary processing of sample	12
2.5 Extraction of DNA	13
2.6 PCR.....	13
2.6.1 PCR reaction system	13
2.6.2 PCR program	13
2.6.3 PCR primer pairs used for amplification of target fragments of 16S rRNA gene	14
2.7 Fluorescence reaction.....	15
3. Results.....	15
3.1 Establishment of an AsCas12a-Based DNA Detection System	15
3.1.1 Expression and purification of Cas12a proteins	15
3.1.2 crRNA preparation and DNA synthesis.....	16
3.1.3 Selection of Target Markers	17
3.2. Collection and Preparation of Water Samples	19
3.2.1 Selection of Sampling Sites.....	19

3.2.2 Processing and Concentration of Water Samples.....	22
3.2.3 Extraction of bacterial genome DNA.....	24
3.3 Detection of Bacterial DNA in Water Samples Using AsCas12a	25
3.3.1 PCR Amplification of Bacterial DNA from Water Samples	25
3.3.2 Cas12a-Based Fluorescence Detection of PCR Products.....	27
3.3.3Detection of the Second-Round Water Samples	29
4. Discussion.....	30
4.1 Development and Application of a CRISPR-Cas12a Assay for Urban Environmental Waters.....	30
4.2 Consistency Between Signal and Site Characteristics.....	31
4.4 Sample Processing for Urban Water Samples.....	31
4.6 Limitations and Future Improvements.....	32
5. Conclusion.....	33
List of instruments	33
References	34

1. Introduction

1.1 Detecting Bacteria in Water Is Essential

Bacteria and bacteriophages are abundant and widely distributed in natural environments and they are important part of aquatic ecosystems (Whitman et al., 1998; Newton et al., 2011). Their abundance and composition vary with water conditions, nutrient availability, and human and animal activities. Detecting bacteria in water is essential to understand the microbial community and changes in the local environment. It is also important for water quality monitoring and for identifying bacteria associated with human or animal sources. Human and animal activity near bodies of water leaves a detectable bacterial trace. In this study, we compared the signals of certain bacteria that indicates nearby human and animal activity. We developed a CRISPR-Cas12a-based method which was applied to three different environments: a high-school campus, a historical landmark, and a local park. The high school campus has a relatively fixed group of people interacting with the environment, and the physical isolation of waters may contribute to a unique composition. *Yuanmingyuan* has a more dynamic interaction between the environment and people due to the changing number and origin of visitors and the connection to outside water systems. Similar to *Yuanmingyuan*, the local park has a large number of visitors and connection to the outside, but with more direct interactions such as shallow shores for duck-feeding.*needs confirmation Apart from direct contact, water infiltration in the soil may indicate activity that occurs over a wider area. Possible forms include farming and underground pipe leakages. All three sites had birds present.

1.3 DNA detection by CRISPR-Cas12a

CRISPR-Cas12a is a nuclease derived from bacteria. It is most widely used in clinical diagnostics, specifically in fields such as infectious disease detection, cancer biomarker detection, genotyping and SNP detection. Cas12a uses a crRNA to recognize a complementary target DNA sequence. Cas12a scans DNA for a T-rich PAM (protospacer adjacent motif) sequence, and it unwinds the DNA if such PAM is present. Then, its crRNA guide attempts to base pair with the unwinded strand. If the base-pairing is successful, the entire enzyme is activated specifically in the RuvC domain responsible for cleavage. In its natural functions, the RuvC domain conducts cis-cleavage (Chen et al., 2018), cutting the target DNA. However, in artificial settings, RuvC conducts trans-cleavage as well (Zetsche et al., 2015), cutting any ssDNA in contact regardless of any traits. This non-specified cleaving is exploited for signal amplification, where a short strand of ssDNA, with a fluorophore and quencher attached to each

end, is used as an indicator to magnify the activation of Cas12a that is otherwise undetectable.) When intact, the ssDNA's short length allows the quencher to inhibit the fluorescent signal of the fluorophore; but after the cut by the RuvC domain of the activated Cas12a, the signal is no longer inhibited and thus detectable. The Cas12a's programmable crRNA guide allows for detection of any target sequence, hence we explored its application in detecting genomes that indicate human and animal activity near waters. A schematic overview of the Cas12a-based detection mechanism is shown in Figure 1_1.

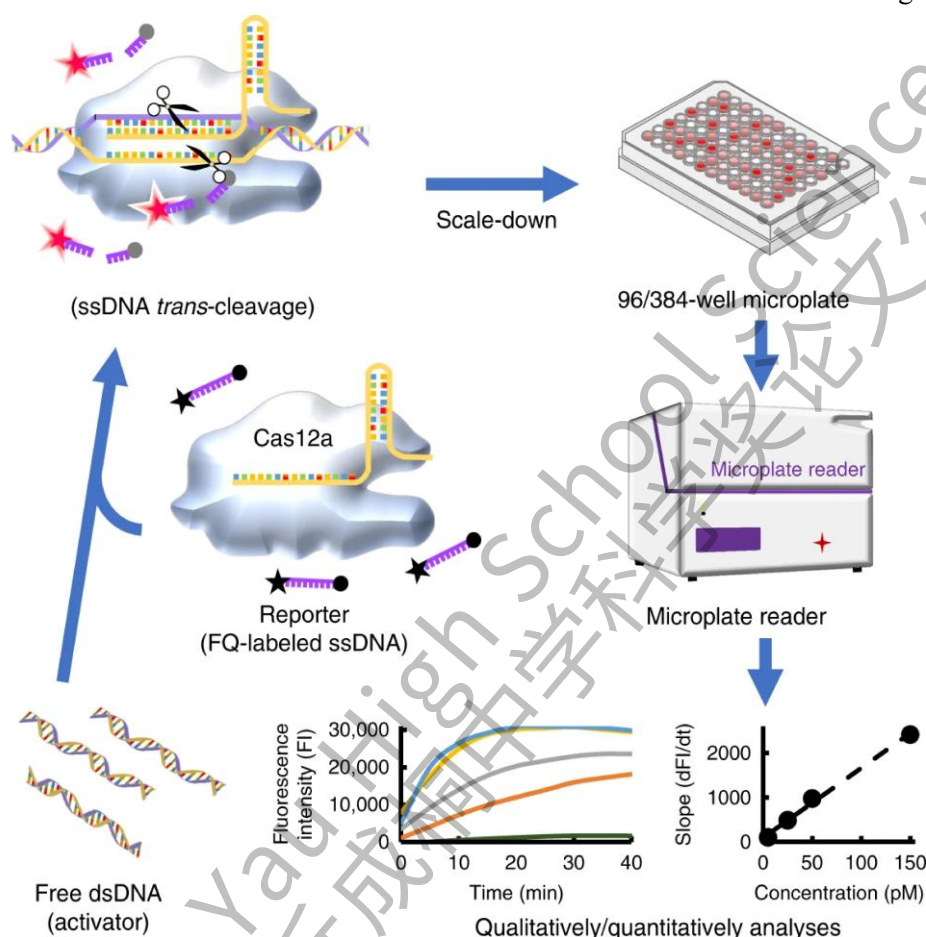


Figure 1_1. The principle of Cas12a nucleic acid detection (Adapted from Liang et al., 2019)

If the extracted double-stranded DNA contains the target DNA harboring the correct PAM and protospacer complementary to the guide of crRNA, it can serve as activators to activate the ssDNA transcleavage activities of Cas12a bound to its cognate crRNA. Consequently, the fluo-reporter will be cleaved. The fluor cannot be quenched by the quencher anymore, and the fluorescen intensity can be read by a microplate reader. Based on this principle, we can detect the specific DNA from environmental samples.

1.4 Nucleic acid extraction and processing

The extraction of nucleic acids is a mature and common procedure, and thus pre-made

extraction kits are available for our use. Taking the kit that we used for example, the extraction of DNA first requires the lysing of cells and the denaturing of proteins, especially DNase; then the washing of residual proteins, salts, clarify , and other impurities involving ethanol-based buffers; and lastly the air-drying of ethanol to prevent disturbance of later steps.

1.5 What We Did in This Study

As students from Beijing 101 Middle School, we wanted to test whether CRISPR-based nucleic acid detection could also be used to study the water environments around us. We collected water samples from Beijing 101 Middle School, *Yuanmingyuan*, and the Olympic Forest Park. We expressed and purified AsCas12a, selected bacterial DNA markers associated with human and animal sources, and tested these markers in water samples from the three sites. Our results show that the Cas12a-based method can detect selected bacterial DNA markers in urban environmental water samples.

2. Methodology

2.1 Cloning, expression and purification of Cas12a proteins

2.1.1 Cloning of Cas12a genes into pET28a vectors

Full-length ORFs encoding *Acidaminococcus sp.* Cas12a protein was synthesized from Sangon Biotech and were cloned into the expression vector containing a His₆-sumo tag at the N-terminus through Gibson Assembly. The protein expression plasmids are kanamycin resistant. The corrected plasmid verified by sequencing was transformed into *E. coli* BL21 (DE3) competent cells.

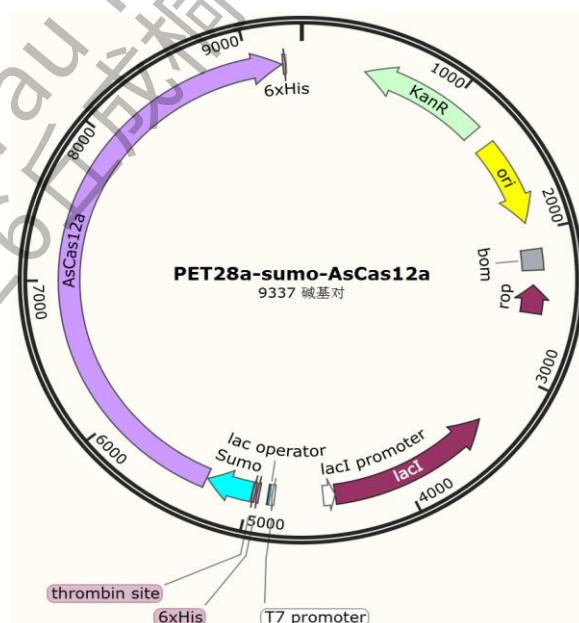


Figure 2_1. The plasmid constructed to overexpress AsCas12a protein

2.1.2 Overexpression of Cas12a proteins

After incubation at 37°C for around 12 h, one colony was inoculated to 80 ml LB medium and cultured for 12 h, and then scaled up to 750 mL LB medium. All media contain 50 µg/ml Kanamycin. When the bacteria grown to OD600 =0.6-0.8, a final concentration of 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to induce protein expression at 16°C for 12 h.

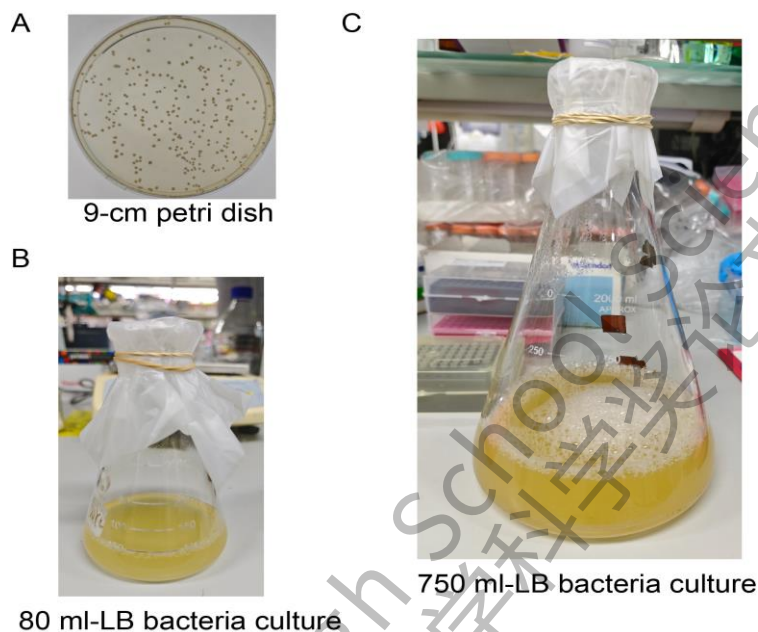


Figure 2_2. The colony and media of BL21 (DE3) cells overexpressing Cas12a proteins

2.1.3 Purification of Cas12a proteins

Cells were harvested by a centrifuge at a speed 4,000 rpm for 20 min, and then resuspended in lysis buffer containing 20 mM Tris-HCl (pH 7.5) and 0.5 M NaCl. Subsequently, cells were disrupted by ultrasonic cell disruptor in ice-water mixture with a power 200 W. The program is 3s on and 8s off for 200 times. Then, the samples were processed with high speed centrifugation 18000 rpm for 1 h at 4°C to separate supernatant from cell pellet.

First Nickel column

Afterward, the supernatant was loaded onto a nickel gravity column containing ~8 mL of Ni²⁺-Sepharose resin (GE Healthcare), and the flow-through (FT) was collected. The resin was then washed and eluted with lysis buffer using an increasing imidazole gradient: 20 mM (200 mL), 50 mM (200 mL), 70 mM (40 mL), 250 mM (40 mL), and 500 mM (20 mL).

Dialysis, His-sumo tag cleavage and second nickel column

The fractions with high purity and amount were selected for dialysis and His₆-sumo tag cleavage. The His₆-sumo tags were cleaved from AsCas12a proteins by His-tagged ubiquitin-like protein 1 (Ulp1) protease during dialysis against buffer containing 20 mM Tris-HCl (pH 7.5), 0.2 M NaCl for 2h at 4°C.

Cas12a protein was further purified by a second nickel column to remove the His₆-Sumo tag. Two columns with ~2 ml resin were used for each Cas12a proteins. The imidazole gradient wash for each column was as follows: flow-through (FT); 50 mM imidazole (20 mL); 300 mM imidazole (20 mL).

Heparin column

Next, the Flow-through fractions of AsCas12a were loaded into Heparin columns for further purification. Buffer A (20 mM Tris pH7.5, 0.2 M NaCl) and Buffer B (20 mM Tris pH7.5, 1.2 M NaCl) were used to elute Cas12a protein bound to the column. The gradient wash program is 0 to 100%B in 30 min, with 2mL/min flow rate. A 40%B gradient was set in the middle. In the end, all fractions collected were verified by SDS-PAGE and stained with Coomassie Blue R250. The target fractions were pooled and then concentrated to a high concentration. The final proteins were aliquoted and flash-frozen to -80°C.

2.2 crRNA preparation and DNA synthesis

crRNA preparation and DNA synthesis

(1) Oligo DNA template preparation for crRNA in vitro transcription

The templates for crRNA in vitro transcription was amplified using two primers, which contains a 25-30 bp overlap part. First, anneal the two primers (20 µM each) by incubation at 95°C for 5 min, and then cooled down with a decrease rate -0.1°C/min to 30°C. Second, mix 5 µL annealed product and 5 µl Phanta Mix DNA polymerase. Incubation at 72°C for 5 min reaction as the template for next step. Next, set up two 50 µl reaction system:

F primer (20 µM) 12 µL

R primer (20 µM) 12 µL

template 1 µL

2*Phanta Mix 25 µL

PCR program:

Initial denaturation at 95°C for 3 min

30 cycles of:

Denaturation at 95°C for 15 s

Annealing at 60°C for 15 s

Extension at 72°C for 30 s

Final extension at 72°C for 3 min

(2) In vitro transcription and purification of crRNA.

The crRNAs were synthesized by in vitro transcription using a T7 RNA polymerase and PCR amplified oligo double-stranded DNA templates. Transcription reactions were performed at 37°C for 6 h with 100 mM HEPES-KOH (pH 7.9), 30mM MgCl₂, 30 mM DTT, 3 mM each NTP, 2 mM spermidine, 0.1 mg/mL T7 RNA polymerase, and 10 ng/μL linearized plasmid DNA template. The RNAs were then purified by gel electrophoresis on a 16% denaturing (8 M urea) polyacrylamide gel and the parts containing the RNA were excised. The excised gel was smashed and soaked in soaking buffer containing 0.38 M NaAc pH5.2, 0.8 mM EDTA pH8.0, 0.2% SDS) for 12 h at 4°C. Thereafter, the RNA solution was subjected to ethanol precipitation. Finally, RNA was resuspended in RNA refolding buffer (50 mM KCl, 5 mM MgCl₂) and refolding using a thermocycler as follows: Incubation at 75°C for 5 min, and the cooled down to 30°C with a decrease rate -0.1°C/min, and stay at 70, 65, 60, 55, 50, 45, 40, 35 and 30°C for 1 min.

(3) Preparation of DNA

All short DNAs used in this study were purchased from TsingKe Biotech. All DNAs were dissolved in buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM MgCl₂. For double-stranded DNA, equimolar amounts of two complementary strands were pre-denatured at 95°C for 10 min and then annealed at room temperature. The Fluo-reporter (5' -6FAM-TTATT-BHQ-1-3') was purchased from Genscript Biotech. It is a single-stranded DNA (ssDNA)

probe with the 5' labeled with 6FAM fluorophore and the 3' labeled with BHQ-1 (Black Hole Quencher 1) quencher. The fluorescence of 6FAM is quenched by BHQ-1 when the reporter is intact. In contrast, after the ssDNA was cleaved, BHQ-1 cannot quench 6FAM anymore, and the fluorescence signal will increase.

Target DNA sequences are listed in the following:

NTS-1B: TAGGAGTTTGGACCGTGTCTCAGTTCCAATGTGGGGGACCTTCC

TS-1B: GGAAGGTCCCCCACATTGGAAGTGAACACGGTCCAACTCCTA

NTS-2C: AGGTATTTTCCGGTAGACGATGGGGATGCGTTCCATTAGATAGT

TS-2C: ACTATCTAATGGAACGCATCCCCATCGTCTACCGGAAAATACCT

NTS-5B: ATCGCTTTTGGATGGACCCGCGGCGTATTAAGTGTGGTAGGG

TS-5B: CCCTACCAACTAGTTAATACGCCGCGGGTCCATCCAAAAGCGAT

DNA primers used to prepare the templates of crRNA 1B, 2C and 5B:

1B-F:

GACTTTAATACGACTCACTATAGGAATTTCTACTCTTGTAGATGACCGTG

1B-R:

AAAAAACCACATTGGAAGTGAACACGGTCATCTACAAGAGTAGAAATTC

2C-F:

GACTTTAATACGACTCACTATAGGAATTTCTACTCTTGTAGATCGGTAGA

2C-R:

AAAAAAGGAACGCATCCCCATCGTCTACCGATCTACAAGAGTAGAAATTC

5B-F:

GACTTTAATACGACTCACTATAGGAATTTCTACTCTTGTAGATGATGGAC

5B-R:

AAAAAATAGTTAATACGCCGCGGGTCCATCATCTACAAGAGTAGAAATTC

2.3 Water sampling

All extractions took place at around 0.5-1m from any solid surface for walking (insert photos).

The water is stirred and scooped from around 50cm below the surface to avoid collection of

large particles. The samples were left still for organizing and transportation for a maximum time of 90 minutes before undergoing preliminary processing. Each sample is around 1L.

2.4 Preliminary processing of sample

Two methods of preliminary processing were proposed. The first method was to use a 10- μ m filtration to remove large inorganic particles, algae, and other multicellular organisms. The liquid then goes through a 0.22 μ m (filtration complex, where bacteria is trapped in the filtering membrane and dissolved into water via repetitive rinsing with 40 mL of ddH₂O.) The second method was to use a 4000rpm centrifuge for 10 minutes and 25 minutes consecutively to ensure that the bacteria is settled along with other particles in the pellet. The supernatant is discarded. 40 mL of ddH₂O is added and stirred with the pellet before another round of centrifuge with 300 rpm for 5 minutes. Bacteria remains in the supernatant while other heavier particles are condensed in the pellet. The supernatant is extracted. The two methods both produce a remaining liquid around 40 mL that goes under centrifuge with 4000 rpm for 15 minutes to force the bacteria into the pellet. The supernatant is removed and the pellet goes to the next procedure.

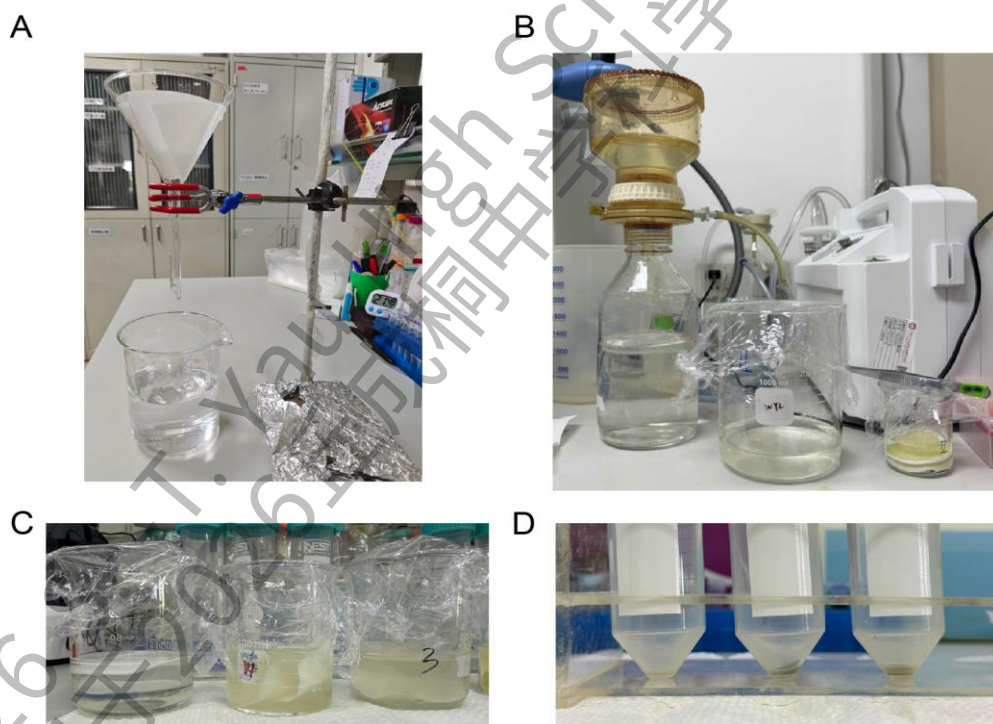


Figure 2_3. The water preprocessing devices.

2.5 Extraction of DNA

For the extraction of bacterial DNA, we used a Bacteria DNA Kit produced by Tiangen Biotech. We followed the provided instructions (place a flowchart here for better representation), except for the last step, applying TE (elution) buffer into the spin column. This is because the kit is (used for longer-term storage of extracted DNA, with the EDTA in the TE buffer inhibiting DNase binding with magnesium ions.) Our next procedure PCR requires DNA polymerase to function properly with magnesium ions, thus we replaced the TE buffer with ddH₂O.

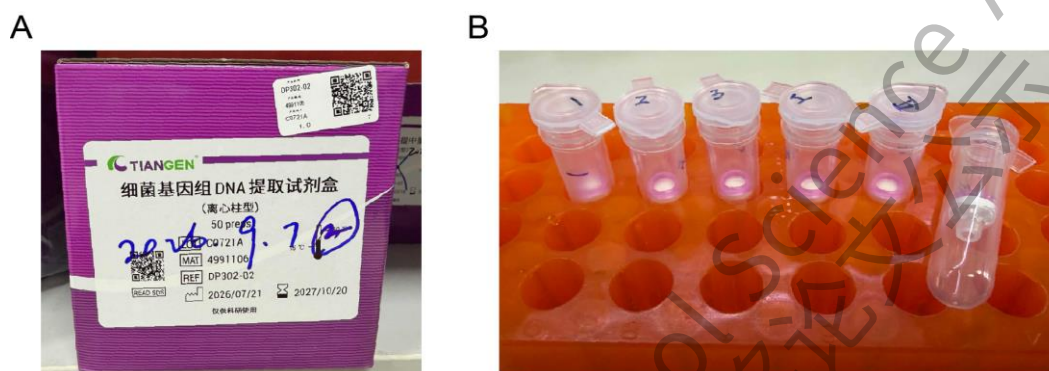


Figure 2_4. The kit used for genomic DNA extraction.

2.6 PCR

2.6.1 PCR reaction system

To amplify the target fragments, Polymerase Chain Reaction (PCR) was performed using 2*Phanta Master Mix (Vazyme) DNA polymerase. Two primer pairs, 1F/1R, 2F/2R and 5F/5R, were used for the amplification of each sample. The PCR reaction was carried out in a 20 μ L total volume in PCR tubes. The 20 μ L reaction system is as follows:

- 2*Phanta Master Mix 10 μ L
- F primer (10 μ M) 1 μ L
- R primer (10 μ M) 1 μ L
- genomic DNA template 50 ng
- addition of ddH₂O to a total volume of 20 μ L

2.6.2 PCR program

The PCR thermal cycling conditions were programmed as follows (Figure 2_5):

Initial denaturation at 95°C for 3 min

40 cycles of:

Denaturation at 95°C for 15 s

Annealing at 60°C for 15 s

Extension at 72°C for 30 s

Final extension at 72°C for 3 min

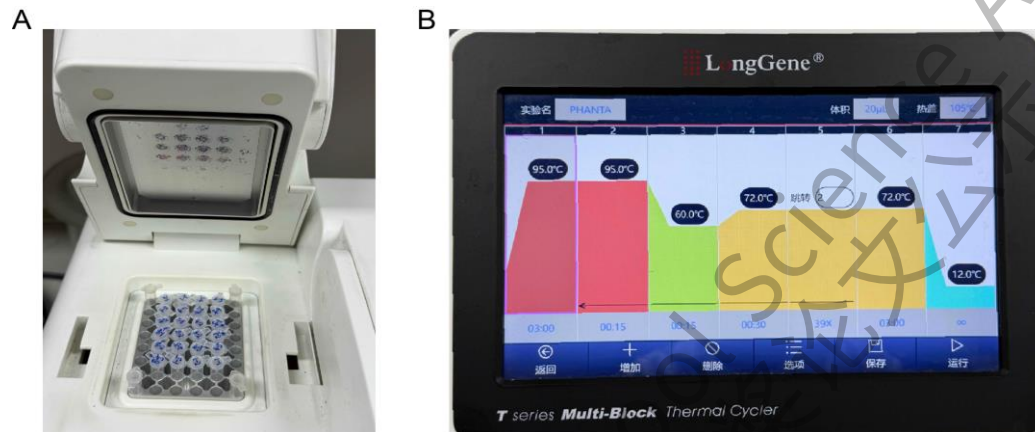


Figure 2_5. The inside and the display panel of the thermocycler

2.6.3 PCR primer pairs used for amplification of target fragments of 16S rRNA gene

(1) GenBac3:

1F: GGGGTTCTGAGAGGAAGGT

1R: CCGTCATCCTTCACGCTACT

(2) HF183:

2F: ATCATGAGTTCACATGTCCG

2R: CTTCTCTCAGAACCCCTATCC

(3) CP_064:

3F: CAGAAGTACAACTCCTAAAAACGTAGAG

3R: GATGACCAATAAACAAGCCATTAGC

(4) GFD:

4F: TCGGCTGAGCACTCTAGGG

4R: GCGTCTCTTTGTACATCCCA

(5) AV4143:

5F: TGCAAGTCGAACGAGGATTTCT

5R: TCACCTTGGTAGGCCGTTACC

2.7 Fluorescence reaction

The fluorescence detection assay was conducted based on the trans-cleavage activity of the CRISPR-Cas12a system. First, the reaction buffer was prepared, containing 20 mM Tris (pH 7.5), 100 mM NaCl, and 10 mM MgCl₂.

For the Cas12a-crRNA complex formation, 20 μ L of 100 nM Cas12a and 20 μ L of 100 nM crRNA were mixed and incubated at 37°C for 15 minutes. Following this, 20 μ L of 2 μ M fluorescence reporter was added to the mixture and mixed gently.

To trigger the trans-cleavage reaction, the PCR products were diluted 3-fold using the reaction buffer. Subsequently, 20 μ L of these diluted PCR products were added to the prepared 60 μ L Cas12a-crRNA-reporter mixture. The final reaction system was incubated at room temperature for 30 minutes in the dark.

Finally, the reaction products were transferred to a black 96-well plate, and fluorescence signals were measured using a microplate reader (EnSpire, Perkin Elmer). The excitation wavelength was set to 485 nm, and the emission wavelength was 520 nm. The intensity of the fluorescence indicated the cleavage of the ssDNA fluo-reporter, and confirming the presence of target DNA sequences in the detected samples.

3. Results

3.1 Establishment of an AsCas12a-Based DNA Detection System

3.1.1 Expression and purification of Cas12a proteins

The AsCas12a and FnCas12a genes were cloned into pET expression vectors with an N-terminal SUMO tag. After the constructs were confirmed by DNA sequencing, the plasmids were transformed into *E. coli* BL21 cells for protein expression. The cells were grown to an OD₆₀₀ of approximately 0.6, and protein expression was induced by adding IPTG to a final concentration of 0.1 mM. After induction for 12 h at 18°C, the cells were harvested by centrifugation and used for subsequent protein purification.

The harvested cells were resuspended in lysis buffer and lysed. After removing the cell debris by centrifugation, the soluble proteins were purified by Ni²⁺ affinity chromatography twice followed by heparin chromatography. The Cas12a proteins were eluted from the first Ni²⁺ affinity column with buffer containing 250 mM imidazole (Figure 3_1A and B). The fractions

containing Cas12a were collected, and ULP protease was added to remove the N-terminal SUMO tag. The proteins were then subjected to further purification by the second Ni²⁺ affinity column to remove the Sumo tag. Most proteins flew through the second nickel column, and some were washed in the wash buffer containing 50 mM imidazole.

Subsequently, both the FT and 50 mM imidazole wash fractions were loaded into a heparin column for further purification. AsCas12a proteins can bind to the heparin column, and no protein flows through the column. When increasing the concentration of sodium chloride, the main peak of AsCas12a was eluted around a conduction value around 52.46 mS/cm, which is corresponding to 42%B buffer B. SDS-PAGE shows that Fractions (15-31) are of high purity . Therefore, fractions (15-31) were pooled and concentrated with concentrator of 50 kDa cutoff (Figure 3_1C and D).

After these three purification steps, both Cas12a proteins showed high purity and good sample homogeneity and were used for subsequent experiments. Finally, the concentration of Cas12a proteins is about 15.4 mg/mL. Proteins were aliquoted to 30 μ L/tube, and flash-frozen to -80°C for use.

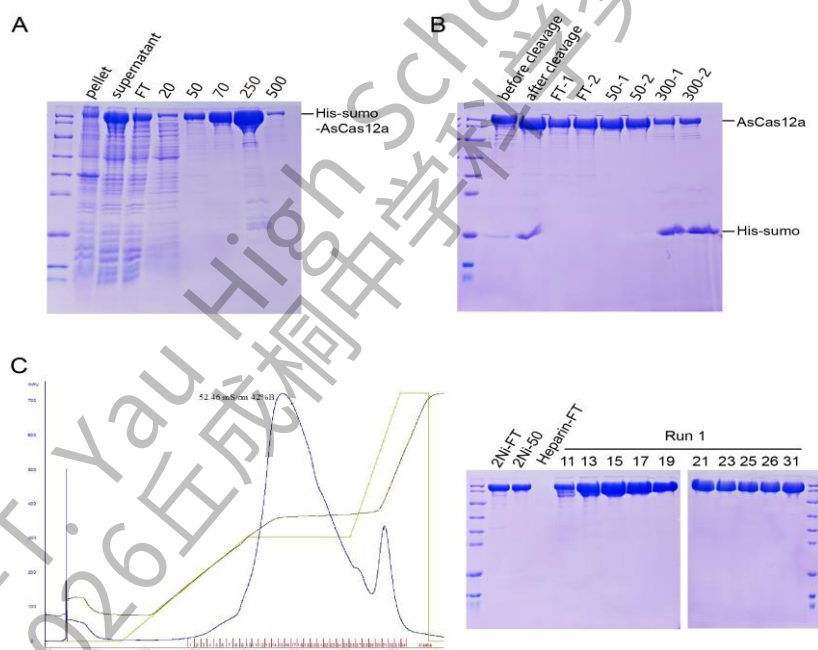


Figure 3_1. Purification of AsCas12a proteins. A. First nickel column; B. Second nickel column; C. Heparin column purification.

3.1.2 crRNA preparation and DNA synthesis

Two oligonucleotides were designed and synthesized based on the target sequence and the sequence requirements for AsCas12a crRNA. The oligonucleotides were used to generate DNA templates by PCR. After amplification, the DNA template for the in vitro transcription was

obtained inferred from the agarose gel, which is longer than the F primer control (Figure 3_2A). DNA template for crRNA transcription contained a T7 promoter at the 5'-prime, and terminated with a continuous poly-T facilitating transcription termination. The final crRNAs for 1B, 2C and 5B are of 45 nt lengths (Figure 3_2B). The first two nucleotides of each crRNAs are 5'-GG which can promote T7 transcription. Finally, Urea-PAGE shows that the three types of crRNAs are prepared correctly (Figure 3_2C).

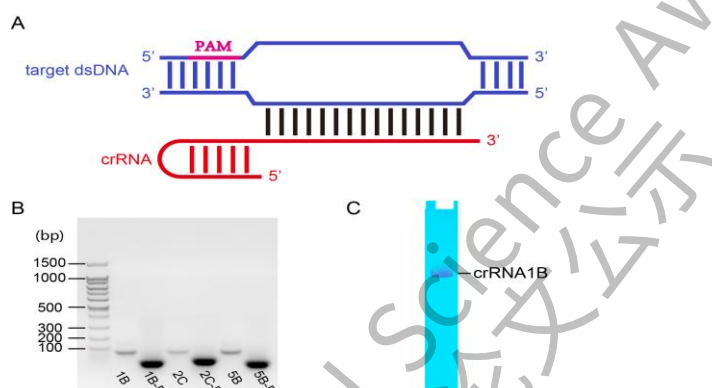


Figure 3_2. Preparation of crRNA and dsDNAs.

(A) Schemes of crRNA and target double-stranded DNA. (B) PCR amplification of DNA templates for crRNA in vitro transcription. (C) Urea-PAGE of a representative crRNA1B.

3.1.3 Selection of Target Markers

We selected five microbial markers that have previously been used in qPCR assays for the detection of fecal-associated microorganisms in environmental water: GenBac3, HF183/BacR287, CPQ-056, GFD, and Av4143. These markers represent different microbial sources, including a general fecal-associated bacterial marker, a human-associated bacterial marker, a human-associated bacteriophage marker, and two avian-associated bacterial markers.

GenBac3 is a general Bacteroidales marker used to detect fecal-associated bacteria from different animal sources (Siefring et al., 2008; Shanks et al., 2010). HF183/BacR287 targets human-associated Bacteroidales, including *Bacteroides dorei* (Green et al., 2014; Haugland et al., 2010). CPQ-056 targets crAssphage, a bacteriophage commonly associated with the human gut (Stachler et al., 2017). GFD and Av4143 are avian-associated markers targeting bacterial groups related to *Helicobacter* spp. and *Lactobacillus* spp., respectively (Rytönen et al., 2021; Shen et al., 2013; Liang et al., 2021; Ohad et al., 2016; Green et al., 2012).

For each marker, we selected a target sequence suitable for AsCas12a recognition and

designed the corresponding crRNA. A synthetic DNA containing a known AsCas12a target was used as the positive control, while a non-target DNA sequence was used as the negative control. The markers and Cas12a targets used in this study are listed in Table 3_1.

Target	Marker	What it represents
General gut-associated bacteria	GenBac3	General Bacteroidales-associated signal
Human gut-associated bacteria	HF183/BacR287	Human-associated <i>Bacteroides/Bacteroidales</i> marker
Human-associated phage	crAssphage (e.g. CPQ_056)	Phage marker strongly associated with human fecal material
Avian-associated bacteria	GFD	Avian-associated <i>Helicobacter</i> marker
Second avian marker	AV4143	Specific avian/bird-associated signal
Positive control	N.A.	Target dsDNA that is complementary to the guide region and harbors the PAM
Negative control	N.A.	Pure water

Table 3_1. The markers and Cas12a targets used in this study.

The published primer pairs for the five markers were used for PCR amplification, as described in Methods Section 2.6.3. The expected amplicon sizes for GenBac3, HF183/BacR287, CPQ-056, GFD, and Av4143 were 129, 126, 126, 123, and 245 bp, respectively. For each PCR amplicon, we searched for a suitable AsCas12a target containing a 5'-TTTV PAM and selected the adjacent protospacer sequence for crRNA design. Each crRNA was designed with a guide sequence complementary to its target DNA.

Before testing environmental water samples, we first checked whether the AsCas12a detection system worked with synthetic target DNA as a positive control. crRNA1B and crRNA5B were selected for the initial test, and their corresponding double-stranded DNA targets were synthesized. Both target DNAs activated reporter cleavage by their respective AsCas12a–crRNA complexes and produced clear fluorescence signals. The raw fluorescence measurements are shown in Table 3_2, and the corresponding data are presented as a bar graph in Figure 3-3. These results confirmed that the designed crRNAs could recognize their target sequences and that the AsCas12a reporter assay was working under our experimental conditions.

crRNA1B	5 min	10 min	15 min	20 min	25 min	30 min	35 min	40 min
target dsDNA	1576	1878	2314	2850	3315	3792	4314	4671
Blank	269	248	292	273	269	273	272	297
crRNA5B	5 min	10 min	15 min	20 min	25 min	30 min	35 min	40 min
target dsDNA	852	1067	1322	1624	1935	2277	2707	2940
Blank	266	272	315	256	285	291	266	284

Table 3_2. The rawdata of target dsDNA activating AsCas12a bound to crRNA1B and crRNA5B in a time course.

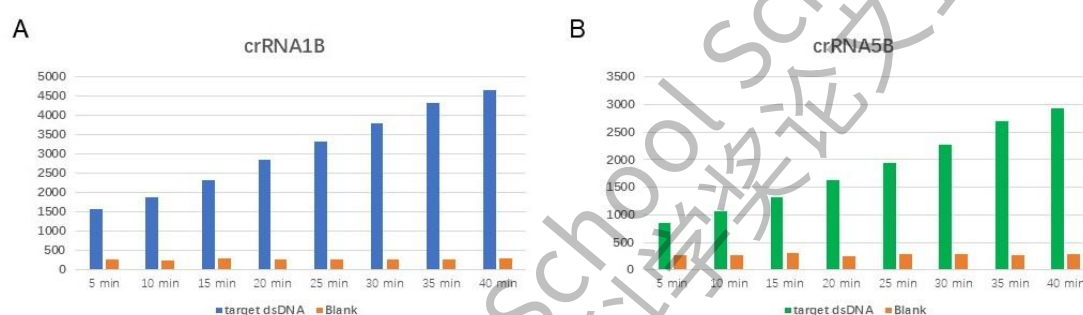


Figure 3_3. Target dsDNA can activate the AsCas12a-crRNA to cleave fluo-reporter robustly

3.2. Collection and Preparation of Water Samples

3.2.1 Selection of Sampling Sites

Water samples were collected from *Yuanmingyuan*, Beijing 101 Middle School, and the Olympic Forest Park. These sites differ in their surroundings and patterns of human and animal activity. *Yuanmingyuan* has extensive park waters with abundant vegetation and waterfowl, together with frequent visitor activity. The campus ponds at Beijing 101 Middle School are closed to *Yuanmingyuan* and exposed to regular student activity and are also home to birds and other animals. The Olympic Forest Park has larger, more open water bodies surrounded by urban green space. Sampling from these three sites allowed us to test the bacterial DNA detection method using water from different environmental settings.

Water sampling was carried out in two rounds. In the first round, two samples (A and B) were collected from Beijing 101 Middle School and used to compare methods for collecting and concentrating bacteria. At the same time, five samples were collected from *Yuanmingyuan* and designated Y1–Y5. After the first round, we adjusted the sampling plan and carried out a second

round of sampling. Three samples were collected from different sites at Beijing 101 Middle School (sampil S1-S3), and three additional samples were collected from Yuanmingyuan and designated Y6–Y8.

For the first sampling round in *Yuanmingyuan*, the five sites differ in human and animal activity as well as in their local water conditions (Figure 3_4). Sites 1 and 4 are isolated water bodies, whereas site2, 3, and 5 are connected. Samples collected from these five sites were designated Y1-Y5, respectively.

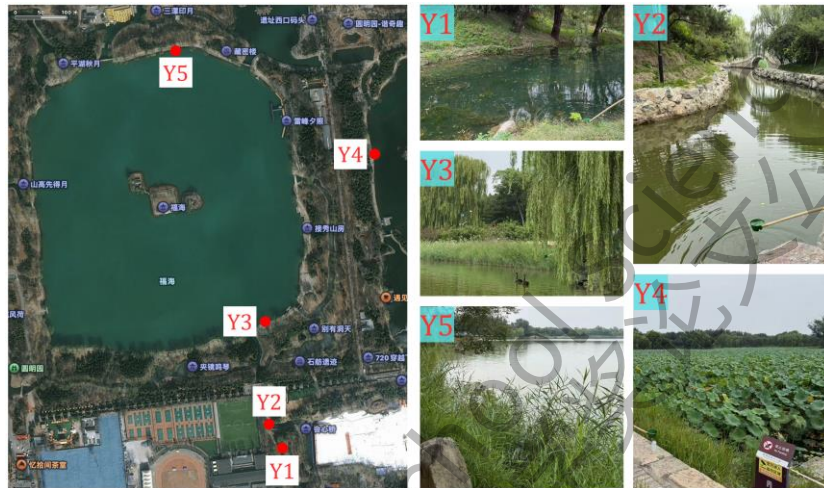


Figure 3_4. Sites of first sampling round in *Yuanmingyuan*

For the second sampling round in *Yuanmingyuan*, the three sites were designated Y6-Y8 (Figure 3_5). All sampling sites are in the same water system. Bird activity was observed in or near the sites across all samples.



Figure 3_5. Sites of second sampling round in *Yuanmingyuan*

For the first sampling round at 101 Middle school, two sites were designated A and B. Both sites belong to a body of water with a purification system, large masses of lotus plants in the center, dwelling of birds, and activities of humans. For the second sampling round at Beijing

101 Middle School, water samples were collected from three separate ponds, with one sample taken from each pond. These samples were designated S1 , S2, and S3.



Figure 3_6. Sites of sampling round in 101 Middle School

S1 was collected from a pond next to a main road in the northern side of the campus where there is no bird activity. S2 and S3 were sampled from the south side of the campus where birds including geese and ducks are active. The two samples came from separate bodies of water. Several ducks and geese were present at site 2, a purification system and a waterwheel were also observed at this site. S3 was collected from a pond with tall and well-grown lotus plants, close to students' outdoor study tables and activity space, where students also eat near the shore. The water at this site was relatively still.

Four sampling sites were selected from different areas of the park. Water samples collected from these sites were designated O1–O4. O1 and O2 are connected, and O1 is downstream of O2. There are many ducks in O1 and the water is turbid and smells not good. The water from O2 is relatively clear, and there are some waterfowls in addition to ducks. There are many tourists along the river of O1 and O2. O3 is far away from O1 and O2, and there are some ducks, and the water is rich in green algae. O4 is a little pond, and the water to be some rain water. No any fowls were observed.



Figure 3_7. Water samples from the Olympic Forest Park.

3.2.2 Processing and Concentration of Water Samples

Optimization of Water Sample Processing

Each water sample normally had a volume of 1 L and contained sediment and floating debris. Some bacteria may also have been associated with these particles, making bacterial collection more difficult. We therefore compared two methods for concentrating bacteria from the water samples.

For this comparison, 2 L of water was collected from each of two sites at Beijing 101 Middle School, designated samples A and B. Each sample was divided into two 1L portions. One portion was processed by Method 1 and the other by Method 2, giving four samples: 1A and 1B for Method 1, and 2A and 2B for Method 2.

Method 1 used two consecutive filtration steps. The water was first passed through a coarse

filter to remove sediment and floating debris. The filtrate was then passed through a 0.22 μm membrane to collect and concentrate most bacterial cells for DNA extraction. We found that agitating the samples before filtration increased bacterial DNA recovery. Coarse filtration also reduced clogging of the 0.22 μm membrane and shortened the time needed to process the samples.

Method 2 used differential centrifugation. The samples were first centrifuged at low speed to remove sediment and larger particles, while most bacteria remained in the supernatant. The supernatant was then collected and centrifuged at a higher speed to pellet the bacterial cells. The resulting pellets were used for DNA extraction.

We compared the two methods based on the amount of DNA recovered and the intensity of the DNA bands. Samples 1A and 1B processed by Method 1 gave stronger DNA bands than samples 2A and 2B processed by Method 2 (Figure 3-8), indicating a higher DNA recovery with Method 1. Two bands were observed. The lower band was consistent with genomic DNA. The upper band was consistently present across both methods and may represent residual high-molecular-weight DNA or material retained at the well; its identity was not determined and it did not affect downstream PCR, as the target amplicons were of the expected sizes.

Based on these results, Method 1 was used to process the remaining water samples. An additional washing step was included before DNA extraction to reduce impurities in the collected material.

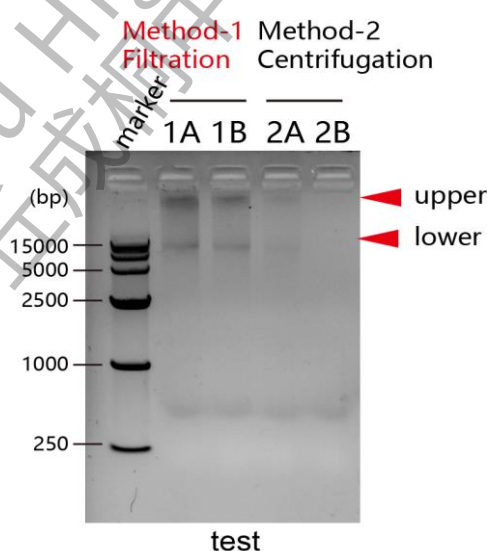


Figure 3_8. The genome DNA extracted from the tested water samples processed with two methods

3.2.3 Extraction of bacterial genome DNA

Bacteria were collected from each water sample using the filtration method described above. The amount of bacterial material collected varied among the samples, as shown by the different amounts of bacterial pellets obtained after processing (Figure 3-9). DNA was then extracted from the collected bacteria and analyzed by gel electrophoresis. The DNA obtained from the three sampling locations is shown in Figure 3-10.

Genomic DNA was extracted from bacteria collected from the three sampling locations: Yuanmingyuan, the Olympic Forest Park, and Beijing 101 Middle School. The extracted DNA was analyzed by agarose gel electrophoresis, and a major DNA band was observed in most samples. Sample S2 from Beijing 101 Middle School, and sample Y5 and Y7 from Yuanmingyuan showed stronger DNA bands, suggesting that more bacterial DNA was recovered from these samples. Interestingly, sample S2 from campus was collected from relatively clear water that had passed through a filtration system. However, ducks and geese were commonly present around this sampling site. The relatively strong DNA band observed for S2 may therefore be related to the presence of these birds and the bacterial input associated with their activity.

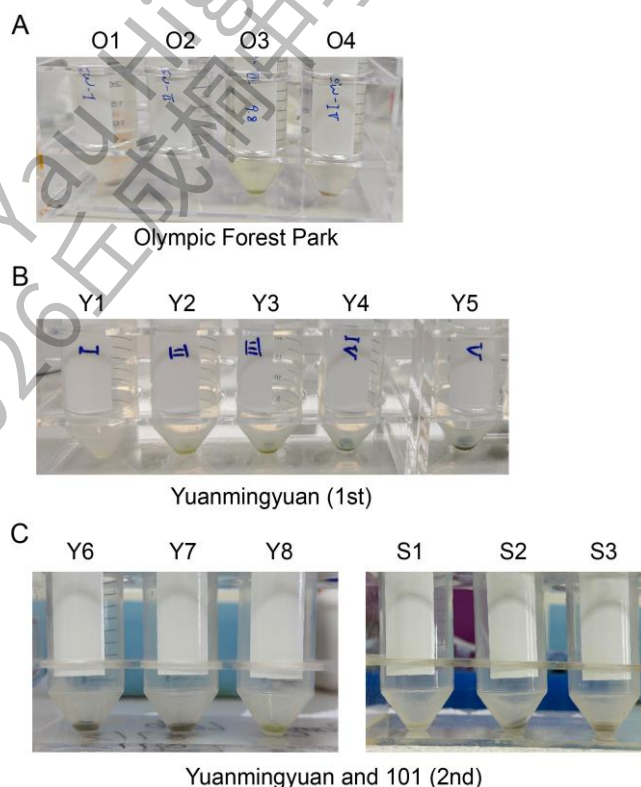


Figure 3_9. Bacteria cell pellets after water preprocessing.

For the three batches of water samples, we used method 1 to obtain the bacterial cells (Figure 3_9). Clear water leads to relatively less cell, and turbid water gives more cells. Some water contains dirt and green algae. However, it will not interfere with the DNA extraction.

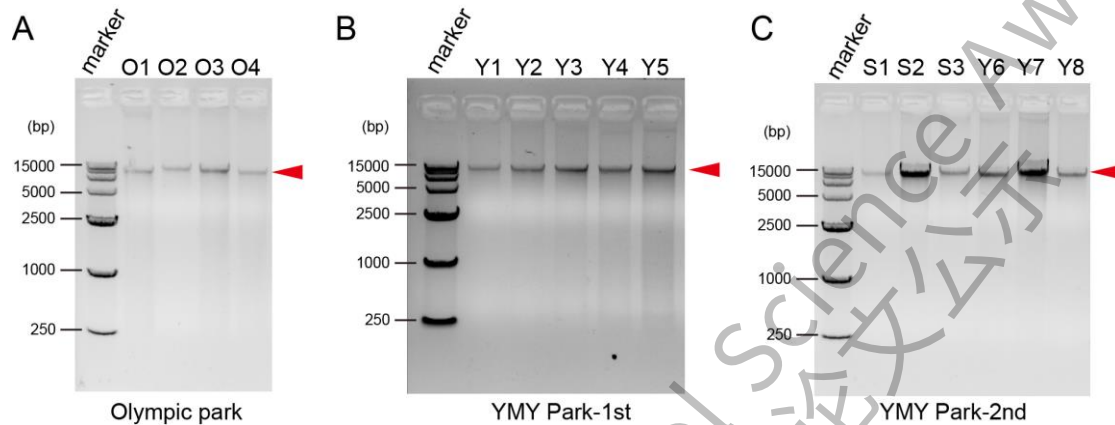


Figure 3_10. Genome DNA extracted from the bacteria cells after water preprocessing.

Using the commercial genome DNA extraction kit, we obtained the DNA from the three batches of water samples. Agarose gel showed that the yields of them differ very much (Figure 3_10). However, the purity is high, and the suspected band with protein contamination observed in Figure 3_8 (upper band) is missing. This indicates that we obtained high-quality genome DNA.

3.3 Detection of Bacterial DNA in Water Samples Using AsCas12a

3.3.1 PCR Amplification of Bacterial DNA from Water Samples

To further test whether our method could be used with real environmental water samples, we first analyzed water samples collected from the Olympic Forest Park. We first tested samples from *Olympic parks* (Figure 3_11). Primer 1 can amplify all the samples. Primer 3 cannot amplify any samples, maybe due to the missing of the phage genome during the extraction of genome DNA. In addition, both primer 4 and 5 can amplify all the samples, and they both detect domestic fowls and water fowl, therefore we only chose primer 5 for further PCR of other water samples.

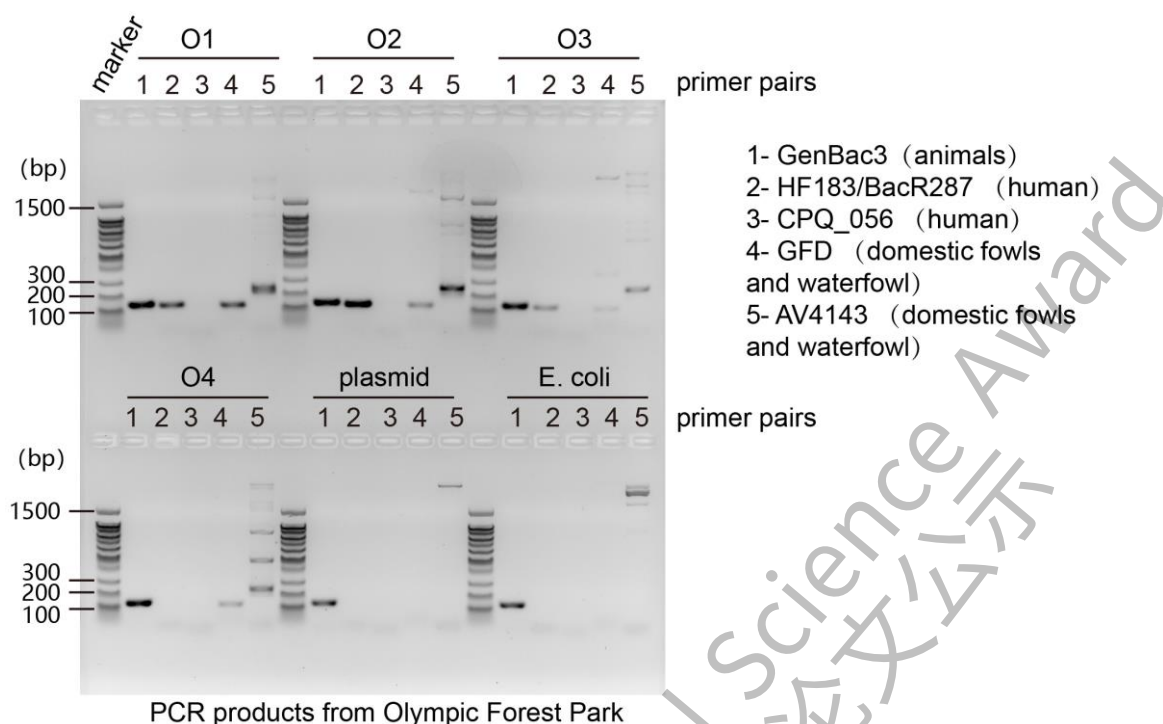


Figure 3_11. PCR amplifications using genomic DNA templates from Beijing Olympic Forest Park.

Next, we extended our analysis to water samples collected from Yuanmingyuan and 101 Middle School (Figure 3_12). For the two batches of Yuanmingyuan samples, primers 1, 2, and 5 were used to amplify the target fragments. Primer 1 amplified the expected 129-bp target fragment from all samples, including the control water and plasmid DNA, indicating the broad detection range of this primer pair. In contrast, only samples Y5, S2 and S3 showed strong bands with primer 2, while almost no signal was detected in the water and plasmid controls, showing the higher specificity of this primer pair. For primer 5, the band intensities varied among the different water samples.

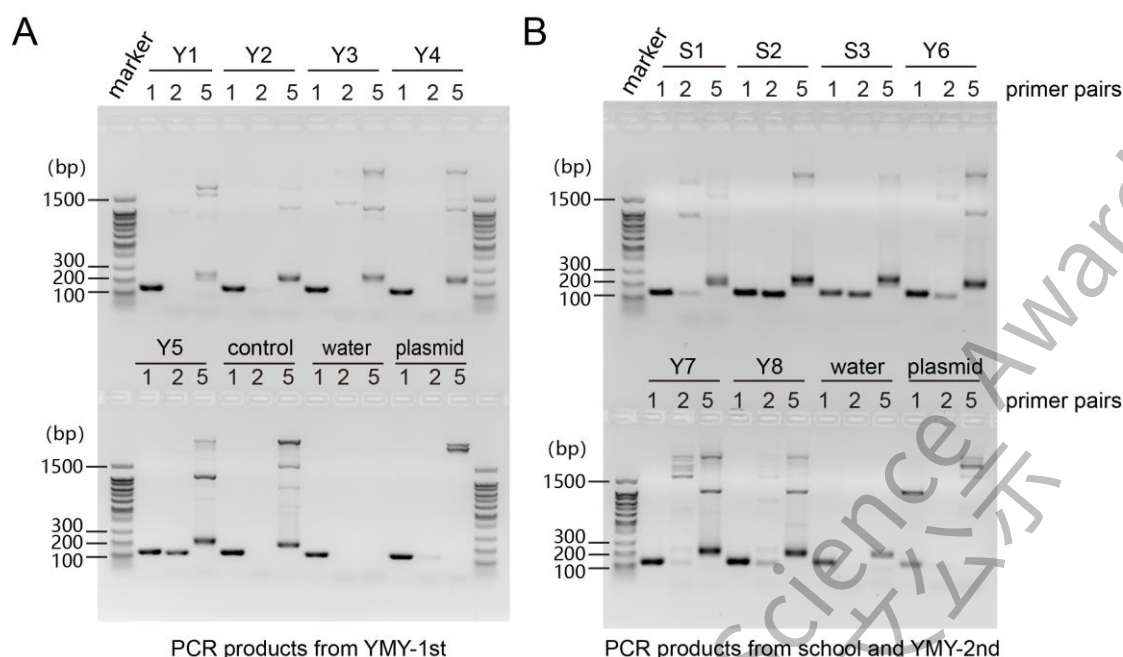


Figure 3_12. PCR amplifications using genomic DNA templates from two batches samples of Yuanmingyuan Park and 101 high school.

3.3.2 Cas12a-Based Fluorescence Detection of PCR Products

Cas12a, crRNA, and the fluorescent reporter DNA were first mixed and incubated for 30 min. The PCR products were diluted threefold and then added to the Cas12a–crRNA–reporter mixture. The reactions were incubated at 37°C for 20 min, and the fluorescence signals were then measured using a microplate reader.

(1) Detection of Olympic Forest Park Water Samples

Water samples O1–O4 collected from the Olympic Forest Park were analyzed using the Cas12a-based fluorescence assay. The fluorescence measurements are shown in Table 3_3, and the corresponding results are presented as a bar graph in Figure 3_13. Samples O1–O4 all can activate Cas12a, showing the presence of the target sequences in the PCR products, indicating the presence of animals. Interestingly, for crRNA2C which target PCR products from primers 2, no signal was detected for sample 4, indicating there is no much human pollution in the little pond, showing in Figure 3_6. Signals were observed for all four samples for crRNA5B targeting PCR products from primers 5, showing the activities of birds and domestic fowls.

All four samples from the Olympic Forest Park showed strong GenBac3 signals, indicating the presence of gut-associated bacteria. Samples O1–O3 also showed strong signals for the human-associated marker, while avian-associated signals were detected in all four samples. The Olympic Forest Park has a more natural environment, with extensive vegetation, abundant bird

activity, and many visitors. The detection of both human- and avian-associated markers is therefore consistent with the surroundings of the sampling sites.

RFU	O1	O2	O3	O4	plasmid	E. coli	target dsDNA	blank
crRNA1B	34819	37404	39675	41905	33262	35338	67065	418
crRNA2C	16863	17478	15817	506	559	663	32877	640
crRNA5B	9532	7389	9653	8383	529	586	28099	458

Table 3_3. Rawdata of the fluorescence signals of AsCas12a activated by samples from Olympic Foreset Park

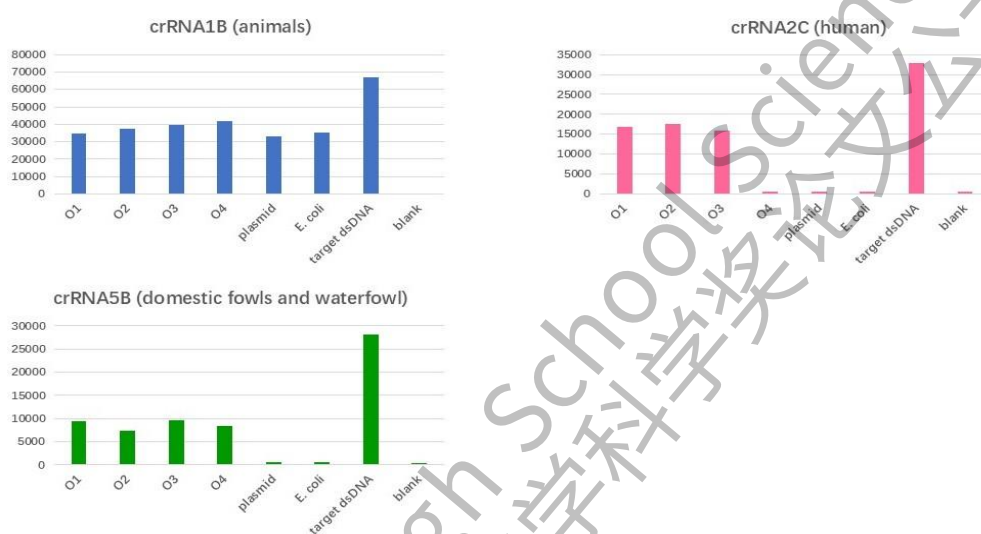


Figure 3_13. Detection of the fluorescence signals of water samples from Olympic Forest Park

(2) Detection of the First-Round Yuanmingyuan Water Samples

We next analyzed the first batch of water samples collected from Yuanmingyuan (Y1–Y5) using the Cas12a-based fluorescence assay. The fluorescence measurements are shown in Table 3_4, and the corresponding results are presented in Figure 3_14. In the first batch of Yuanmingyuan samples (Y1–Y5), the general gut-associated bacterial marker GenBac3 produced relatively strong signals in most samples. In contrast, a strong signal for the human-associated marker HF183 was mainly observed in Y5. The avian-associated markers showed relatively weak signals across all five samples.

Interestingly, Y1 showed a different pattern from the other samples. This sample was collected from a relatively isolated site with few visitors, and the water was turbid with a noticeable unpleasant odor. Y1 showed a strong GenBac3 signal, but relatively weak signals for both the human- and avian-associated markers. This suggests that the gut-associated bacteria detected in Y1 may have come from sources not well represented by the human and avian markers used

in

our

study.

RFU	Y1	Y2	Y3	Y4	Y5	water	plasmid	target dsDNA	blank
crRNA1B	44950	49693	39455	43720	47845	51902	52648	87013	588
crRNA2C	641	4771	573	638	17858	564	9389	43127	744
crRNA5B	1019	2811	4577	2213	3601	444	535	47298	632

Table 3_4. Raw data of the fluorescence signals of AsCas12a activated by samples from the first-round Yuanmingyuan Water Samples.

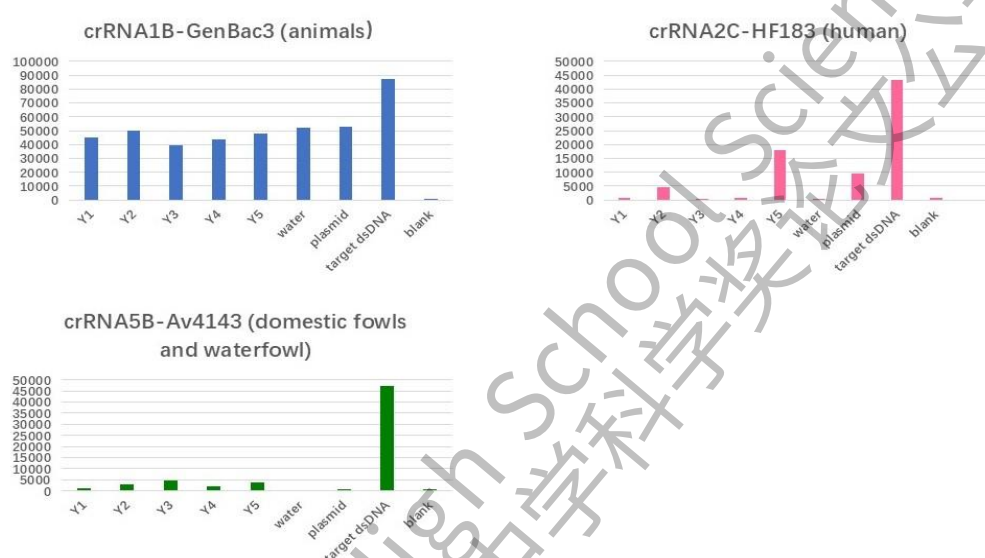


Figure 3_14. Detection of the fluorescence signals of water samples from first-round Yuanmingyuan Water Samples

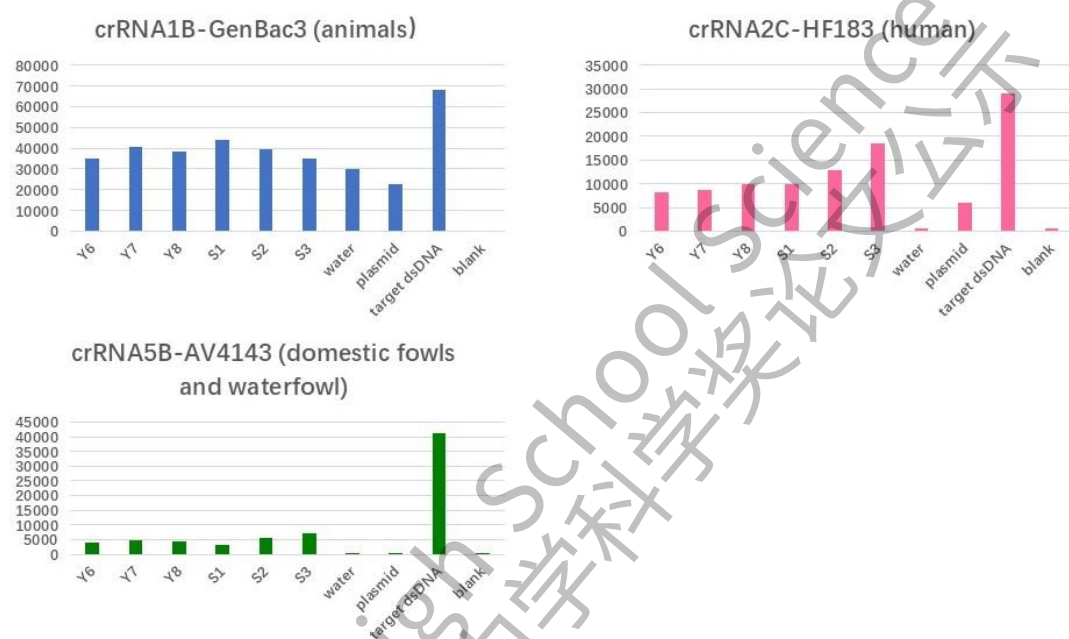
3.3.3 Detection of the Second-Round Water Samples

We next analyzed the second batch of water samples collected from Yuanmingyuan (Y6–Y8), together with three samples collected from Beijing 101 Middle School (S1–S3), using the Cas12a-based fluorescence assay. The fluorescence measurements are shown in Table 3_5, and the corresponding results are presented in Figure 3_15.

In the second batch of Yuanmingyuan samples and three samples from 101 Middle School, GenBac3, a general gut-associated bacterial marker, produced strong signals in most samples. The human-associated marker HF183 also showed relatively strong signals. In contrast, the avian-associated markers produced only weak signals. These results indicate that general and human-associated gut bacterial markers were more readily detected in these samples than the avian-associated markers.

RFU	Y6	Y7	Y8	S1	S2	S3	water	plasmid	target dsDNA	blank
crRNA1B	34820	40710	38278	44200	39735	35033	30204	22574	67957	356
crRNA2C	8144	8792	9871	9892	12758	18570	633	6109	29164	704
crRNA5B	4104	4833	4459	3260	5623	7205	556	582	41348	389

Table 3_5. Rawdata of the fluorescence signals of AsCas12a activated by samples from the second-round Yuanmingyuan Water Samples



Data analysis of the 2nd water samples from Yuanmingyuan Park

Figure 3_15. Detection of the fluorescence signals of water samples from second-round Yuanmingyuan Water Samples

4. Discussion

4.1 Development and Application of a CRISPR-Cas12a Assay for Urban Environmental Waters

In this study, we developed a technique for detecting bacterial DNA in urban environmental waters using CRISPR-Cas12a. We cloned, expressed, and purified AsCas12a and selected specific regions of the 16S rRNA gene for PCR amplification and Cas12a detection. We then applied the assay to water samples collected from three urban environments. Bacterial DNA was detected in most samples, including avian gut-associated bacterial signals in several samples. These results show that Cas12a can be used to detect specific bacterial DNA markers in urban environmental waters.

4.2 Consistency Between Signal and Site Characteristics

Fluorescent signals of crRNA1B are of strong magnitude across all samples because the six sampling sites are all under the influence of animals. Signals of crRNA2C are stronger in the three 101 middle school samples than in the three *Yuanmingyuan* samples. This is due to the higher concentration of human activity in the 101 campus than in *Yuanmingyuan* and the smaller scale of water bodies in the campus. Among the three 101 samples, S3 demonstrated higher signals of crRNA2C than S1. This is because students study and eat near the banks of Y3, while they rarely stay around Y1. crRNA5B signals are similar across the three *Yuanmingyuan* samples due to the direct connection between sites and the even occurrence of bird activity in the water system. crRNA5B signals in S2 and S3 are higher than that of S1 due to the absence of bird activity in and near Y1. Despite bird activity present at Y2 during sampling, its larger mass and the presence of a filtration system might have reduced the signal.

4.3 CRISPR-Cas12a Detection Does Not Require Bacterial Culture

Bacteria are commonly detected by culture-based methods. For example, bacterial culture is widely used in hospitals to identify bacteria from patient samples. However, culture-based methods are relatively slow and may take several days to culture bacteria. In addition, many environmental bacteria are difficult to grow under laboratory conditions (Staley & Konopka, 1985). In contrast, bacterial culture is not required in CRISPR-Cas12a detection. The total time cost is just several hours, much faster than canonical culture-based bacterial detection. These advantages enable the general use of this method for bacterial DNA detection.

4.4 Sample Processing for Urban Water Samples

The two filtration steps served different purposes. The large sediment particles and floating debris were removed during the first coarse filtration step, thus preventing rapid clogging of the 0.22- μm membrane. The second step concentrated bacteria and other microorganisms from the prefiltered water onto the 0.22- μm membrane. Together, these steps made it possible to process larger volumes of water more efficiently and reduced the time required for microbial concentration.

We found that agitating the samples before filtration increased bacterial DNA recovery. Coarse filtration also reduced clogging of the 0.22- μm membrane and made sample processing faster. These observations suggest that some bacteria may be associated with sediment or suspended particles. Agitation before filtration may help release these bacteria into the water and improve

their recovery.

A possible limitation of our experiments is that bacteria that remain tightly associated with large particles may be removed during coarse filtration, resulting in the loss of some bacteria. The main concern is that bacteria associated with large particles may be lost when sediment is removed, whereas insufficient removal of sediment may interfere with filtration and DNA extraction. Further optimization will be needed to balance microbial recovery with filtration efficiency. The combination and further comparison of filtration and centrifugation methods may help improve bacterial recovery while keeping the procedure simple.

4.5 Target Selection Affects Bacterial Detection

The amount of DNA recovered from each sample did not correlate directly with the Cas12a fluorescence signal. Samples with more total DNA did not necessarily produce stronger signals. This is expected because the extracted DNA represents the overall bacterial community, whereas our Cas12a assay detects only the specific bacterial markers selected in this study.

Interestingly, some relatively clear water samples, such as S2 and S4, gave strong positive signals, whereas several turbid or foul-smelling samples showed only weak signals. A weak signal does not necessarily indicate fewer bacteria. Our targets were mainly selected from 16S rRNA gene sequences associated with human and avian gut bacteria and therefore represent only a small part of the bacterial community. Turbid or odorous water may contain other bacteria associated with organic matter and its decomposition that are not detected by these markers. This could explain why some samples contained relatively large amounts of DNA but produced weak Cas12a signals.

These results also show an important limitation of our current assay: the bacteria we can detect depend strongly on the markers we choose. Additional 16S rRNA gene targets would allow a broader range of bacteria to be detected. Sequencing the water samples could first identify the major bacterial groups present and help us choose more suitable Cas12a targets.

This could also extend the method beyond detecting gut-associated bacteria. For example, future work could identify bacterial groups associated with turbid or foul-smelling urban water and design Cas12a assays specifically for these organisms. Such assays could help investigate the microbial changes associated with poor water quality and unpleasant odors.

4.6 Limitations and Future Improvements

The main limitation of this study is the small number of bacterial markers tested. Human- and avian-associated markers cannot represent the full bacterial community in environmental water.

This may explain why some samples with obvious differences in appearance did not show corresponding differences in our assay.

We also tested a phage marker associated with human gut bacteria but did not detect a clear signal. Two avian-associated markers were included, but signals were detected in only a few samples. The bacteria or phages may have been present at levels below our detection limit, or the markers we selected may not have been suitable for these water samples.

16S rRNA gene sequencing could help identify the major bacterial groups in these samples. We could then select new Cas12a targets based on the bacteria actually found in the water, rather than relying only on markers selected in advance. Testing more bacterial and phage markers would also give a better signals of the microorganisms present at different sites.

5. Conclusion

In conclusion, AsCas12a showed good activity and could detect selected microbial DNA in environmental water samples. By using different crRNAs, the method can be adapted to different DNA targets and may be useful for microbial detection in water. Further work is needed to improve sample processing and detection sensitivity and to test more microbial targets. With these improvements, Cas12a-based detection may become a useful method for monitoring microbial DNA in environmental water.

List of instruments

1. Thermo Scientific Nanodrop2000
2. LongGene T Series multi lock Thermal Cycler for PCR
3. PerkinElmer Multimode Plate Reader for fluorescence reaction signal detection
4. Ultrasonic Cell Disruptor
5. Centrifuge
6. Nickel Gravity Column
7. Heparin Column
8. UV Gel Imager (Tanon MINI SPACE 3000)

References

1. Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., & Doudna, J. A. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science*, 360(6387), 436–439. <https://doi.org/10.1126/science.aar6245>
2. Green, H. C., Dick, L. K., Gilpin, B., Samadpour, M., & Field, K. G. (2012). Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. *Applied and Environmental Microbiology*, 78(2), 503–510. <https://doi.org/10.1128/AEM.05734-11>
3. Green, H. C., Haugland, R. A., Derminhaci, M. M., Lamendella, R., & Shanks, O. C. (2014). Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. *Applied and Environmental Microbiology*, 80(10), 3086–3091. <https://doi.org/10.1128/AEM.04137-13>
4. Haugland, R. A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., & Shanks, O. C. (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology*, 33(6), 348–357. <https://doi.org/10.1016/j.syapm.2010.06.001>
5. Liang, H. X., Yu, Z. S., Liu, R. Y., Zhang, H. X., & Wu, G. (2021). Development and application of molecular markers on microbial source tracking for avian. *Acta Ecologica Sinica*, 41(3), 1006–1014. <https://doi.org/10.5846/stxb202003210633>
6. Liang, M., Li, Z., Wang, W., Liu, J., Liu, L., Zhu, G., & Huang, J. (2019). A CRISPR-Cas12a-derived biosensing platform for the highly sensitive detection of diverse small molecules. *Nature Communications*, 10(1), Article 3672. <https://doi.org/10.1038/s41467-019-11648-1>
7. Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews*, 75(1), 14–49. <https://doi.org/10.1128/MMBR.00028-10>
8. Ohad, S., Ben-Dor, S., Prilusky, J., Kravitz, V., Dassa, B., Chalifa-Caspi, V., Kashi, Y., & Rorman, E. (2016). The development of a novel qPCR assay-set for identifying fecal contamination originating from domestic fowls and waterfowl in Israel. *Frontiers in Microbiology*, 7, Article 145. <https://doi.org/10.3389/fmicb.2016.00145>
9. Rytönen, A., Tiwari, A., Hokajärvi, A. M., Uusheimo, S., Vepsäläinen, A., Tulonen, T., & Pitkänen, T. (2021). The use of ribosomal RNA as a microbial source tracking target highlights the assay host-specificity requirement in water quality assessments. *Frontiers in Microbiology*, 12, Article 673306. <https://doi.org/10.3389/fmicb.2021.673306>

10. Shanks, O. C., White, K., Kelty, C. A., Hayes, S., Sivaganesan, M., Jenkins, M., Varma, M., & Haugland, R. A. (2010). Performance assessment of PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology*, 76(5), 1359–1366. <https://doi.org/10.1128/AEM.02033-09>
11. Shen, Z., Duan, C., Zhang, C., Carson, A., Xu, D., & Zheng, G. (2013). Using an intervening sequence of *Faecalibacterium* 16S rDNA to identify poultry feces. *Water Research*, 47(16), 6415–6422. <https://doi.org/10.1016/j.watres.2013.08.013>
12. Sieftring, S., Varma, M., Atikovic, E., Wymer, L., & Haugland, R. A. (2008). Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument platforms. *Journal of Water and Health*, 6(2), 225–237. <https://doi.org/10.2166/wh.2008.022>
13. Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., & Shanks, O. C. (2017). Quantitative CrAssphage PCR Assays for Human Fecal Pollution Measurement. *Environmental Science & Technology*, 51(16), 9907–9915. <https://doi.org/10.1021/acs.est.7b02703>
14. Staley, J. T., & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, 39(1), 321–346. <https://doi.org/10.1146/annurev.mi.39.100185.001541>
15. Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, 95(12), 6578–6583. <https://doi.org/10.1073/pnas.95.12.6578>
16. Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Yan, W. X., Samarney, A., Tabebordbar, M., Zhang, F., & Regev, A. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163(3), 759–771. <https://doi.org/10.1016/j.cell.2015.09.038>