



# GLAPD: Whole Genome Based LAMP Primer Design for a Set of Target Genomes

Ben Jia<sup>1†</sup>, Xueling Li<sup>2†</sup>, Wei Liu<sup>2</sup>, Changde Lu<sup>2</sup>, Xiaoting Lu<sup>2</sup>, Liangxiao Ma<sup>2</sup>, Yuan-Yuan Li<sup>2\*</sup> and Chaochun Wei<sup>1,2\*</sup>

<sup>1</sup> Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, <sup>2</sup> Shanghai Center for Bioinformation Technology, Shanghai, China

## OPEN ACCESS

### Edited by:

Abd El-Latif Hesham,  
Assiut University, Egypt

### Reviewed by:

Debmalya Barh,  
Institute of Integrative Omics  
and Applied Biotechnology (IOAB),  
India

Pallavi Singh,  
Northern Illinois University,  
United States

### \*Correspondence:

Yuan-Yuan Li  
yyli@scbt.org  
Chaochun Wei  
cwei@sjtu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Food Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 17 June 2019

**Accepted:** 26 November 2019

**Published:** 13 December 2019

### Citation:

Jia B, Li X, Liu W, Lu C, Lu X,  
Ma L, Li Y-Y and Wei C (2019)  
GLAPD: Whole Genome Based LAMP  
Primer Design for a Set of Target  
Genomes. *Front. Microbiol.* 10:2860.  
doi: 10.3389/fmicb.2019.02860

Loop-mediated isothermal amplification (LAMP) technology has been applied in a wide range of fields such as detection of foodborne bacteria and clinical pathogens due to its simplicity and efficiency. However, existing LAMP primer designing systems require a conserved gene or a short genome region as input, and they can't design group-specific primers. With the growing number of whole genomes available, it is possible to design better primers to target a set of genomes with high specificity based on whole genomes. We present here a whole Genome based LAMP primer designer (GLAPD), a new system to design LAMP primer for a set of target genomes using whole genomes. Candidate single primer regions are identified genome wide and then combined into LAMP primer sets. For a given set of target genomes, only primer sets amplifying them and only these genomes will be output. In order to accelerate the primer designing, a GPU version is provided as well. The effectiveness of primers designed by GLAPD has been assessed for a wide range of foodborne bacteria. GLAPD can be accessed at <http://cgm.sjtu.edu.cn/GLAPD/> or <https://github.com/jiqingxiaoxi/GLAPD.git>. A simple online version is also supplied to help users to learn and test GLAPD: <http://cgm.sjtu.edu.cn/GLAPD/online/>.

**Keywords:** LAMP, group-specific primer, whole genome based primer designer, foodborne pathogens, primer design

## INTRODUCTION

Loop-mediated isothermal amplification (LAMP) is a simple-operating, effective and reliable method to amplify DNA sequence (Notomi et al., 2000; Parida et al., 2008; Mori and Notomi, 2009). The amplification is under a constant temperature (about 62°C) and the running time is short (within 1 h). In many application scenarios, LAMP is a better option than polymerase chain reaction (PCR) because the reaction can be in small and portable devices (Curtis et al., 2012; Chaumpluk et al., 2016). A basic LAMP primer set contains four synthetic primers derived from six primer regions (**Supplementary Figure S1**). Therefore, LAMP primer design is more complex than PCR primers. Designing a LAMP primer set to specifically identify a group of genomes (group-specific) at the same time has a high demand in many application fields like foodborne harmful bacteria detection, clinical pathogen identification, agricultural pathogen identification, and so on.

The group-specific primers mean that they can be applied to many target genomes belonging to a group (the primers are common) and at the same time the primers can't amplify any other genomes not included in this group (the primers are specific) (Jarman, 2004; Kalendar et al., 2017).

For example, there are 16 strains of white spot syndrome virus with complete genomes and more than 100 thousand other viruses in NCBI nucleotide database (up to October 29th, 2018). A group-specific primer for white spot syndrome virus should only amplify the 16 strains but no other viruses. Traditionally, the group-specific primers are designed based on conserved genes (Peng et al., 2015), genome regions (Yao et al., 2016) or the multiple sequence alignment (MSA) of these genes or genomic regions (Kurosaki et al., 2010). But this method is limited by the small number of suitable genes and the difficulty to generate MSA for a large number of sequences (Chen and Tompa, 2010). In addition, the primers based on traditional methods often can't meet the requirements in practice. For example, the LAMP primer set from Wang et al. (2015) targeting *Staphylococcus aureus* was not able to amplify some *S. aureus* strains and might amplify some unexpected genomes (more details in results part). With more whole genomes available (O'Leary et al., 2016), it's a better method to design group-specific primers based on whole genomes (Treven, 2015; Demkin et al., 2017).

There are some systems for LAMP primer design now. The most popular one is PrimerExplorer V5<sup>1</sup>, an online software. However, the maximum length of its input sequence is limited up to 2,000 bps. Therefore, designing primers based on the whole genome is not provided. In addition, common or specific primers can be designed by PrimerExplorer V5 using MSA results by clicking the "Common" or "Specific" button separately. However, these two buttons can't work at the same time, which makes it not straightforward to design group-specific primers. Another system is LAVA (Torres et al., 2011), which can design common primers for a group of target genomes. However, it requires MSA results as the input, which limits the target region to a gene or conserved genomic region. LAVA doesn't check the specificity of primers. FastPCR (Kalendar et al., 2017) is a system that can design LAMP primers using the whole genome. Similar as PrimerExplorer, it has an online version, and it designs common primers or specific primers separately. It can't design group-specific primers in one run. According to the authors' best knowledge, no existing system can design group-specific LAMP primers using whole genomes.

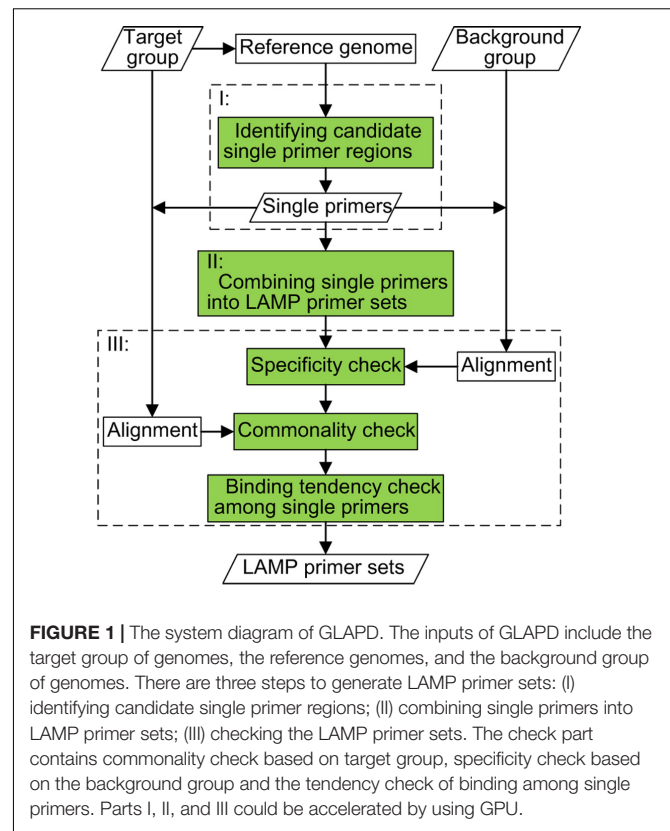
Here we present GLAPD (whole genome based LAMP primer designer), a new system to design group-specific LAMP primer sets. By using the whole genome sequences as input data, GLAPD can ensure the specificity of the primers and increase the chance to design a successful primer set. A graphics processing unit (GPU) version of GLAPD is also provided.

## MATERIALS AND METHODS

### The Genome Based LAMP Primer Design System GLAPD

The system diagram of GLAPD is listed in **Figure 1**. GLAPD has three steps: (I) identifying candidate single primer regions; (II) combining single primers into LAMP primer set; and (III) checking the LAMP primer set. The inputs, outputs and the computation steps are listed below in more details.

<sup>1</sup><http://primerexplorer.jp/e/>



### Inputs

In this step, two groups are defined first. The target group is defined as a group of genomes or genome regions which are expected to be amplified by the LAMP primer set. The primers generated by GLAPD are expected to identify each target genome. If it failed to generate a primer set for the group of target genomes, the system will output primers that can amplify the maximum number of genomes.

Similar to the target group, a background group is defined as a group of genomes or genome regions which are not expected to be amplified. Primers designed by GLAPD should not amplify any genome in the background group.

One genome from the target group needs to be selected as the reference genome which will be used as the template to generate primer sequences. The reference genome can be randomly picked from the target group.

### Identifying Single Primer Regions

A basic LAMP primer set contains four synthetic primers from six primer regions, named F3, F2, F1c, B1c, B2, and B3. Sequences from F1c and F2 are synthesized into one primer FIP and sequences from B1c and B2 are synthesized into another primer BIP. The positions relationship among these single primers are showed in **Supplementary Figure S1**. In order to design a LAMP primer set, those candidate primer regions are identified first. They are then combined into LAMP primer sets.

Genome based LAMP primer designer identifies all candidate primer regions in the reference genome according to primer's

length, GC-content, melting temperature ( $T_m$ ), stability and so on (**Supplementary Table S1**). The secondary structure of primer is checked by GLAPD using the thermodynamical approach similar to Primer3 program (Untergasser et al., 2012). At the same time, the end of each primer is checked to exclude symmetric sequences and homopolymers. GLAPD uses customized parameters to identify primer regions according to the GC-content of the target region (between F3 and B3). If GC-content of the target region is high, the GC-content and  $T_m$  of primers are set to be high, vice versa.

### Combining Single Primers Into Primer Sets

Primers from six regions are combined into one basic LAMP primer set. GLAPD uses the positional relationship (**Supplementary Figure S1**) among the six regions, GC-content relationship between primers and whole amplification region (**Supplementary Table S1**),  $T_m$  relationship among primers (the  $T_m$ s of F1c and B1c are 3°C higher than other primers) to combine them. Then the combined LAMP primer set will be checked for commonality, specificity and tendency of binding among single primers.

### Checking Commonality and Specificity of the LAMP Primer Set

An ideal LAMP primer set should be able to amplify all the target genomes but not genomes from the background group. In order to do this check, firstly, all single primers are aligned to the target genomes and the background genomes using Bowtie (Langmead et al., 2009). By default, no mismatch is allowed when a primer is aligned to the target group. If GLAPD fails to design LAMP primer sets to amplify all target genomes, a small number of mismatches are allowed when primers are aligned to the target group. However, if a primer can be aligned to a background genome within two mismatches (by default), this primer is considered as not specific. The more mismatches required to align the primer to the background genomes, the more specific is the primer. No matter how many mismatches in a primer, no mismatch is allowed in the 5' of F1c and B1c, and the 3' of F3, F2, B2, B3, LF, and LB primer. After the alignment, primers' positions, strand information (plus or minus) and the number of mismatches in each genome are recorded.

Using the information generated above and the positional relationship of the six primer regions (**Supplementary Figure S1**), GLAPD checks the ability of a LAMP primer set to amplify genomes in the background group. If a LAMP primer set can amplify any background genome, this set will be discarded. Therefore, more flexible thresholds for positional relationship among primers can be used to improve the specificity in this step. After the LAMP primer set passes this specificity check, the number of genomes or genome regions in the target group amplified by the primer set is calculated using the same method in specificity check.

### Checking the Binding Tendency of Any Two Primers

The LAMP primer set passed the commonality and specificity check will be checked for every single primer's tendency of binding to other single primers in this primer set.

This check uses the thermodynamical approach similar as the Primer3 program does.

### Outputting LAMP Primer Sets

The LAMP primer set passed all above check steps will be output. When GLAPD has designed 10 (by default) LAMP primer sets successfully, or GLAPD has checked all candidate LAMP primer sets, the system stops automatically. The outputs contain the sequences, positions, lengths of the primers and genomes which can be amplified. The LAMP primer sets are not overlapped with each other. By default, the shortest distance between two LAMP primer sets' F3 regions is 300 bps.

### Loop Primers

In order to accelerate the amplification, two additional loop primers (LF and LB) can be added (**Supplementary Figure S1**). GLAPD can also design LAMP primer set with loop primers. The candidate primer regions are identified for loop primers from the reference genome first. Those candidate regions must meet the requirements listed in **Supplementary Table S1**. Then, GLAPD combines loop primers with other single primers into a LAMP primer set. A LAMP primer set could contain one or two loop primers. The  $T_m$ s of loop primers are set to be 3°C higher than  $T_m$  of F3, F2, B2, and B3. At last, this LAMP primer set is checked for the tendency of primer annealing and its commonality.

### GPU Version

Graphics processing units can be used to accelerate GLAPD in three steps. In the step of identifying candidate single primer regions, GLAPD can identify them from many positions of the reference genome simultaneously. Many primers' GC-content, stability,  $T_m$  and secondary structure can be calculated in parallel in GPUs. In the step of combining single primers, each thread of GPU is assigned with a different candidate F3 primer then GLAPD tries to design all LAMP primer sets containing this F3 primer in parallel. In the checking step, each thread calculates the number of target genomes and background genomes that can be amplified by the primer set designed in this thread, and every single primer's tendency of binding to other single primers. In each thread, only the LAMP primer set amplifying the maximum number of target genomes will be returned to CPU for output.

### Databases

Three databases were generated. Database-1 was the database of complete genome sequences of bacteria and archaea, which was downloaded from NCBI's FTP on August 5th, 2013. It contained 4,902 sequence files from 2,599 strains (about 9GB).

Database-2 was the database of all complete mitochondrion sequences of suina, bovine, and caprine, which was downloaded from NCBI nucleotide database on June 14th, 2016. In this database, 120, 394, and 209 mitochondrion sequences for suina, bovine and caprine, respectively were selected.

Database-3 was the database of complete genome sequences of bacteria and archaea downloaded from NCBI's Nucleotide database on September 19th, 2018. It contained 15,728 bacterial sequence files and 494 archaea sequence files (about 60GB).

Database-4 was the database of complete genome sequences of viruses downloaded from NCBI's Nucleotide database on August 2th, 2017. It contained 163,576 viruses sequence files.

## DNA Extraction

The activated strains were cultured in appropriate method (**Supplementary Table S2**) and collected when the cultures reached an optical density at 600 nm (OD600) between 0.6~1.0. The genomic DNA of the strains were extracted using AxyPrep™ Multisource Genomic DNA Miniprep Kit (Axygen Bioscientific, Inc., United States). The samples were ground into powder with liquid nitrogen and homogenized with 400 µl of cell lysis buffer, and all other steps followed the manufacturer's instructions.

## LAMP Reaction

Loop-mediated isothermal amplification reaction was performed in a 25 µl reaction mixture. The mixture contained 1 × ThermoPol Buffer (contained 2 mmol/L MgSO<sub>4</sub>), 6 mmol/L MgSO<sub>4</sub> (total 8 mmol/L), 1.4 mmol/L of each dNTP, 1.6 µmol/L of each inner primers (FIP and BIP), 0.2µmol/L of each outer primers (F3 and B3), 8 units of Bst DNA Polymerase (Large Fragment).

When test the LAMP primer designed for *S. aureus*, the mixture contained 20 ng of template DNA of each 29 bacterial strains. When test the LAMP primer's specificity designed for *Vibrio vulnificus* and *Vibrio cholerae*, the mixture also contained 20 ng of template DNA of each 29 bacterial strains and when test its commonality, the amount of template DNA was 50 ng, 5 ng, 500 pg, 50 pg, 5 pg, 500 fg, 50 fg, 5 fg and 0.5 fg, respectively. In all NTC (no template control) reaction, template DNA was replaced by sterilized water.

The LAMP reaction was carried out at 62°C for 60 min using a Veriti™ Dx Thermal Cycler (Thermo Fisher, United States), then inactivated Bst DNA Polymerase at 80°C for 10 min. After the reaction, 1 µl 1000× SYBR Green I was added into the solution to confirm whether the reaction occurred. In a positive reaction, the color of the solution was green and in a negative reaction, the color was orange.

## RESULTS

### Experimental Validation of the Group-Specific Primers Designed by GLAPD

#### Group-Specific LAMP Primers of *Staphylococcus aureus*

*Staphylococcus aureus* is one main type of foodborne pathogens around the world (Kadariya et al., 2014; Paudyal et al., 2017). Traditionally, its *nuc*, *mecA* genes were used to design the LAMP primer sets (Wang et al., 2015; Chen et al., 2017). But the two genes are not conserved among all *S. aureus* (Hoegh et al., 2014; Karmakar et al., 2016) and may exist in other *Staphylococcus* spp. (Borjesson et al., 2015). Therefore, existing LAMP primer sets (Wang et al., 2015) for *S. aureus* are neither common nor specific enough (**Supplementary Tables S3, S4**).

In database-1 the *S. aureus* species had 43 strains (**Supplementary Table S3**). GLAPD used those 43 strains as the target group and the rest of the genomes in database-1 as the background group then designed several group-specific LAMP primer sets. One set (**Table 1**) was located in a predicted gene which codes 50S ribosomal protein L14. This LAMP primer set was common for all *S. aureus* and not for any other bacteria. It was validated by experiments (**Supplementary Figure S2**).

#### *Vibrio* Group-Specific LAMP Primer Set

*Vibrio* spp. is one of the main pathogenic bacteria in seafood (Huehn et al., 2014; Mizan et al., 2015). Most LAMP primers are designed only for the identification of *V. cholerae* (Okada et al., 2010) or *V. vulnificus* (Han et al., 2011). In spite that the *V. vulnificus* and *Vibrio parahaemolyticus* can be detected simultaneously in one reaction, it required two LAMP primer sets in the reaction (Wang et al., 2016).

In database-1, there were eight strains of *V. cholerae* and three strains of *V. vulnificus* (**Supplementary Table S5**). *V. cholerae* and *V. vulnificus* belonged to *Vibrio* genus. GLAPD used those 11 strains as the target group and the rest of the genomes in database-1 as the background group. Several group-specific LAMP primer sets were designed by GLAPD. One set (**Table 2**) was located in a gene which coded 30S ribosomal protein S20. This LAMP primer set was common for all *V. cholerae* and *V. vulnificus*, and was specific to these bacteria only. It was validated by experiments listed in **Supplementary Figure S3**.

Besides the two group-specific LAMP primer sets mentioned above, we have applied GLAPD to design group-specific primers for more than ten other foodborne pathogens. GLAPD could successfully design group-specific primers for most of these foodborne pathogens and more than half of those primers worked well in real experiments (**Supplementary Table S6**).

The flexible setting of target and background group.

When GLAPD is used to design group-specific primers, the target group and background group could be defined flexibly by users. Some examples are listed below to show this flexibility.

**TABLE 1** | The LAMP primer set designed by GLAPD for *S. aureus*.

| Primer | Sequence (5'-3')                             |
|--------|--|
| F3     | GTCATTACGACGAACACC                           |
| B3     | ATCCAACAAGAAACACGC                           |
| FIP    | TTAAAAATGCAACACCAGGTGG-CTTAGTACGTACGATTACAGC |
| BIP    | GCTGTTTTACGACCAGATCCA-TGAAAGTAGCAGACAACCTCT  |

**TABLE 2** | The LAMP primer set designed by GLAPD for *V. cholerae* and *V. vulnificus*.

| Primer | Sequence (5'-3')                            |
|--------|---|
| F3     | CGAGACTTGTGACGAGCTG                         |
| B3     | TCCAAGCTGAGAAACGCTCG                        |
| FIP    | TGCAACTGCTGCATTGCGTGTG-TAAGGCCTTTAGTCGCCATG |
| BIP    | CCAGCAGCGATTGCGACGAAC-ACAATGCTAGCCGTCGTTT   |



(1) To design group-specific primer for all genomes in a genus. *Salmonella* is a common bacterial pathogen responsible for salmonellosis, a common disease affects the intestinal trace and it can cause substantial socioeconomic burden (Liu et al., 2018). In database-3, 471 complete genomes and 1 assembly genome with 66 contigs were *Salmonella*. GLAPD could design the genus-level salmonella-specific LAMP primer set successfully. One LAMP primer set designed by GLAPD (**Supplementary Table S7**) was in a gene coding (2E,6E)-farnesyl diphosphate synthase. It could amplify all 472 genomes without any mismatches, and it was specific to all target genomes considering all other genomes in database-3 as the background genomes.

(2) To design group-specific primer for some genomes in a genus. Both *V. cholerae* species and *V. vulnificus* species belong to *Vibrio* genus. In database-3, 44 complete genomes were *V. cholerae* and 19 genomes were *V. vulnificus*. The primer set listed in **Table 2** was still common in all 63 genomes without any mismatches, and is specific to the target genome considering the other genomes in database-3 as the background group (don't allow any mismatches).

(3) To design group-species primer for all genomes in a species. In database-3 there were 351 complete genomes of *S. aureus*, the primer set in **Table 1** was neither common nor specific in the new database. One new primer set (**Supplementary Table S8**) designed based on database-3 was overlap with two genes coding S4 domain-containing protein *YaaA* and DNA replication repair protein *RecF*. It was common for all 351 genomes without any mismatches and specific to the target genomes as well.

(4) To design group-specific primer for some genomes in a species. Enterohemorrhagic *Escherichia coli* O157:H7 was a major foodborne pathogen and it caused diarrhea (Tarr et al., 2005; Lim et al., 2010). The *E. coli* O157:H7 was one group of *E. coli* species. In database-3, there were 59 complete genomes of *E. coli* and three were O157:H7 (In order to exclude ambiguous group, only take care of part *E. coli* genomes, those accession number must start with "NC"). One LAMP primer set (**Supplementary Table S9**) designed by GLAPD overlaps with two genes coding recombinase and peptide transporter. This primer set was common for all three genomes without any mismatch and specific to all these three genomes as well (with all other genomes in database-3 as the background group).

## GLAPD Can Design Primers for Other Organisms

Genome based LAMP primer designer was first developed for rapidly detecting foodborne pathogens using LAMP technology. But group-specific primers for other organisms can also be designed by GLAPD. For example, we have tried to design group-specific primers to detect halal products using GLAPD. Pork is not allowed in halal products (Nakyinsige et al., 2012). The previous LAMP primer sets for pork identification were designed on *DN1* gene and *cytb* gene located in mitochondria (Yang et al., 2014; Ran et al., 2016). But the LAMP primer set (Yang et al., 2014) couldn't cover all mitochondria genomes (**Supplementary Table S10**). In database-2, there were 120

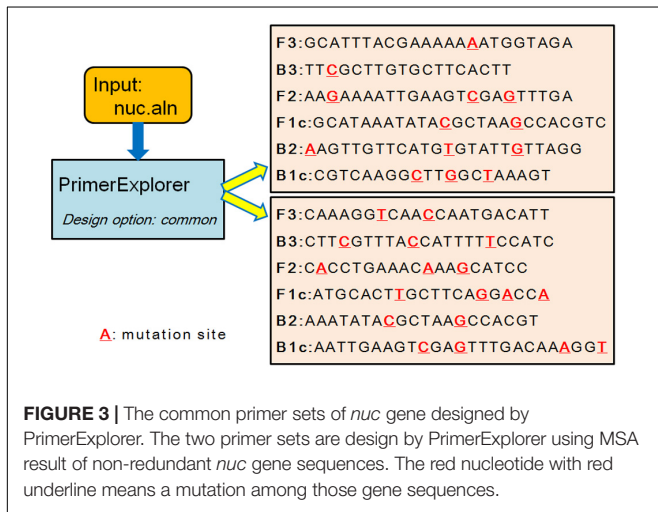
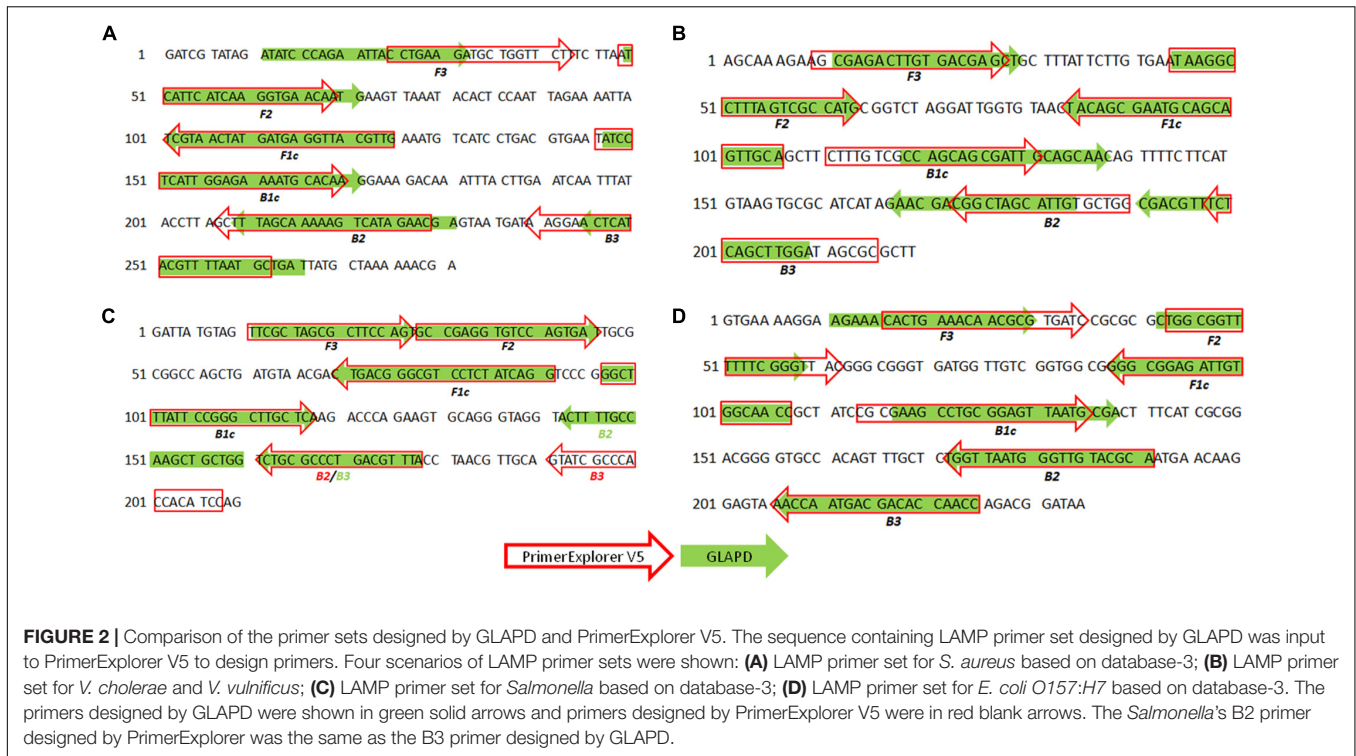
mitochondria sequences of sunia which were considered to be from pork products. We used GLAPD to design several LAMP primer sets with the 120 sunia mitochondrion sequences as the target group and the rest in database-2 as the background group. One of these primer sets (**Supplementary Table S11**) was in a 16S rRNA gene. This LAMP primer set was common for all pork mitochondria and would amplify the pork mitochondria only.

We also tried to design group-specific primers for several types of aquatic animal viruses in database-4. GLAPD could design group-specific primers for most aquatic animal viruses using default parameters (**Supplementary Table S12**). The group-specific primers of infectious hematopoietic necrosis virus and spring viraemia of carp virus were validated by experiments. If GLAPD couldn't design group-specific primers, it would output the most common primers.

## Comparison With Existing Systems

PrimerExplorer V5 was an online LAMP primer design software. A candidate genomic region was required as the input. However, GLAPD did not require the prior knowledge about the candidate gene or genomic region. Using the same sequences from the candidate regions containing primer sets designed by GLAPD, the results of PrimerExplorer V5 and GLAPD were very similar. Most primers from two designers overlapped and some of them were identical (**Figure 2**). The main reason of differences was the different primer combination strategies.

Existing systems such as PrimerExplorer, LAVA and FastPCR can design common primer set for a group of genomes without considering the specificity comparing to a background group (LAVA was not used to design primers because it was not downloadable anymore). Users can input a set of genomic regions considering the first region as the target and the rest as the background group to design specific primers. However, it is not straightforward to design primer sets directly with a given target group and a given background group at the same time. For example, the *mecA* and *nuc* genes were used as candidate regions (Wang et al., 2015; Chen et al., 2017) to design primers for *S. aureus*. In database-3, there were 351 *S. aureus* genomes, among which 235 genomes had *mecA* gene and all 351 genomes had *nuc* gene. The non-redundant 19 *mecA* sequences and 44 *nuc* sequences were used as inputs for PrimerExplorer and FastPCR to design primers. For *mecA* gene, the MSA result was generated by clustalW (Thompson et al., 1994) with default parameter, and only 68.1% nucleotides were identical. PrimerExplorer couldn't design any LAMP primer sets using this MSA result and FastPCR also couldn't design any primers using the 19 *mecA* sequences. For *nuc* gene, the MSA alignment result was also generated by clustalW and only 74.7% nucleotides were identical. FastPCR couldn't design any common primers for the 44 sequences. PrimerExplorer could design two LAMP primer sets. However, because of the low similarity, there were many mutations in the designed primer regions in different target genomes (**Figure 3**, red underline), which indicated that the commonality of this primer set was not good. Therefore, using PrimerExplorer and FastPCR to design common primer set for *S. aureus* with the *mecA* and *nuc* gene regions as input was not successful. On the contrary, GLAPD could use the whole genome to design



LAMP primer sets for the *S. aureus* genomes successfully. More examples are in **Supplementary Material (Supplementary Figure S4 and Supplementary Table S14)**.

More comparisons between GLAPD and those existing systems were listed in **Table 3**.

## DISCUSSION

### Why Do We Use Group-Specific Primers?

Detecting multiple foodborne pathogens simultaneously can help the food safety because it can speed the detection of pathogens.

When use LAMP or PCR, the traditional method for this aim is adding multiple sets of primers into one assay. Each primer set can only be applied to one pathogen. This method has some disadvantages: (1) a large number of different primers in one assay may increase the risk of generating primer dimer, and the efficiency of one primer set could be inhibited by other primer sets (Xu et al., 2012; Zhao et al., 2014); (2) It's a challenge task to find a feasible combination of multiple primer sets due to the huge number of different combinations of candidate primer sets in addition to the test of each primer set for each target genome.

Genome based LAMP primer designer, on the other hand, can design group-specific primers, which can avoid those disadvantages. In one assay, only one primer set is needed. Less primers in assay can decrease the risk of interactions among primers and reduce the test workload as well.

### Designing Primers Based on the Whole Genome

Traditionally, primers are designed for some conserved genes or a small genome region. Most regions of the whole genome are neglected. In addition, many primer design systems have a limited length of input sequences. If no primers could be designed by GLAPD, the chance to find suitable primers would be low.

Genome based LAMP primer designer can use whole genomes as input directly. It scans all candidate single primers derived from the whole genome, then all candidate primers are combined into primer sets and tested one by one. This can vastly improve the success rate of primer design.

**TABLE 3** | The comparison with existing LAMP primer designing systems.

|                     | GLAPD         | PrimerExplorer | LAVA         | FastPCR                             |
|---------------------|---------------|----------------|--------------|-------------------------------------|
| Input               | Sequence      | Sequence/MSA   | Sequence/MSA | Sequence/MSA                        |
| Input length        | Not limited   | ≤2,000 bps     | Not limited  | Not limited                         |
| common primers      | Yes           | Yes            | Yes          | Yes                                 |
| specific primers    | Yes           | Yes            | No           | Yes                                 |
| Group-specific      | Yes           | No             | No           | No                                  |
| Graphical interface | No            | Yes            | No           | Yes                                 |
| Operating system    | Unix/Linux    | Windows        | Unix/Linux   | Windows                             |
| Version             | Stand-alone   | Online         | Stand-alone  | Online/Stand-alone                  |
| Language            | C,CUDA C,Perl | Java           | Perl         | Java                                |
| License             | GPL2          | Free           | BSD          | Online: free<br>Stand-alone: charge |

## A New Strategy Is Used to Design Group-Specific Primers

The group-specific primers can be designed based on conserved genes, genomic regions or MSA results. However, existing methods have limitations: (1) There are limited number of well-known conserved genes in each organism. This number will be much smaller for a slightly larger number of different target organisms. (2) The conserved genes may exist in some background organisms. For example, the 16S rRNA genes are conserved in many bacteria, therefore non-target organisms may also be amplified if they contain very similar 16S rRNA genes; (3) It is difficult to generate a MSA results from a big number of input sequences, and it is almost impossible to generate MSA for many genomes.

Genome based LAMP primer designer uses a different strategy to avoid those problems. Firstly, GLAPD searches all candidate primers genome wide, then the candidate single primers are aligned with target and background genomes. The alignment information about positions and strands is used to check primers' commonality and specificity. The distance between two single primers can be different in each target genome as long as the distance is within the allowed range. The sequences between two primers can also be different in different target genomes as long as the primer regions are conserved among target genomes. In other words, GLAPD can design group-specific primers in variable genome regions with a higher success rate. This strategy can also be used in designing primers for other amplifications, like PCR.

## The Parameter Configuration in GLAPD

Parameters and thresholds used in GLAPD were similar with PrimerExplorer V5. Good results were still achieved if GLAPD used loose thresholds. For example, when the 5' stability of F1c or B1c was set to be bigger than -4 kcal/mol, the experimental results of the primers were still good. More experiments might be needed to decide better parameters or thresholds for a specific group of organisms.

## GPU Version of GLAPD

In order to accelerate the LAMP primer design, a GPU version of GLAPD was developed. The GPU version was very promising to accelerate the primer design procedure in identifying candidate

single primer regions (GPU version is three time faster than CPU version) and combining single primers to a primer set may be slower than CPU version in some scenarios (**Supplementary Table S13**). We are currently working on it to improve the performance of the GPU version.

## CONCLUSION

Designing group-specific primers is a difficult task for amplifications like PCR, and it is even more challenging for LAMP due to the number of primers in LAMP primer set. Here we present a new LAMP primer designer, GLAPD, to design a LAMP primer set targeting on a group of genomes. Instead of using well-known gene regions, the whole genome could be used directly for primer design, which increased the success rate.

Genome based LAMP primer designer could be applied to design LAMP primers for the identification of any organisms without known regions as input. The results of GLAPD are similar to PrimerExplorer V5 when the same sequences are input. The effectiveness of GLAPD were validated in experiments. With GLAPD, the chance to successfully design a LAMP primer set to identify a group of organism is higher than before and it can be a good system to accelerating the application of LAMP technology in many fields such as food quarantine, epidemic disease surveillance and so on. GLAPD can be downloaded from <http://cgm.sjtu.edu.cn/GLAPD/> or <https://github.com/jiqingxiaoxi/GLAPD.git>. Users can also learn and test GLAPD using the simple online version: <http://cgm.sjtu.edu.cn/GLAPD/online/>.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

CW: conceptualization and design of the system. BJ and CW: system implementation. BJ and LM: system test. XLi, WL, CL, XLu, and Y-YL: experimental validation of designed LAMP



primers. BJ, XLi, Y-YL, and CW: writing the manuscript. All authors have read and edited the manuscript.

## FUNDING

This work was supported by grants from the National Natural Science Foundation of China (61472246), the National Basic Research Program of China (2013CB956103), the National High-Tech R&D Program (863) (2014AA021502), and Cross-Institute Research Funding of Shanghai Jiao Tong University (YG2017ZD01 and YG2015MS39).

## REFERENCES

- Borjesson, S., Gomez-Sanz, E., Ekstrom, K., Torres, C., and Gronlund, U. (2015). *Staphylococcus pseudintermedius* can be misdiagnosed as *Staphylococcus aureus* in humans with dog bite wounds. *Eur. J. Clin. Microbiol. Infect. Dis.* 34, 839–844. doi: 10.1007/s10096-014-2300-y
- Chaumpluk, P., Plubcharoensook, P., and Prasongsuk, S. (2016). Rapid detection of aflatoxigenic *Aspergillus* sp. in herbal specimens by a simple, bendable, paper-based lab-on-a-chip. *Biotechnol. J.* 11, 768–779. doi: 10.1002/biot.201500435
- Chen, C., Zhao, Q., Guo, J., Li, Y., and Chen, Q. (2017). Identification of Methicillin-resistant *Staphylococcus aureus* (MRSA) using simultaneous detection of *mecA*, *nuc*, and *femB* by loop-mediated isothermal amplification (LAMP). *Curr. Microbiol.* 74, 965–971. doi: 10.1007/s00284-017-1274-2
- Chen, X., and Tompa, M. (2010). Comparative assessment of methods for aligning multiple genome sequences. *Nat. Biotechnol.* 28, 567–572. doi: 10.1038/nbt.1637
- Curtis, K. A., Rudolph, D. L., Nejad, I., Singleton, J., Beddoe, A., Weigl, B., et al. (2012). Isothermal amplification using a chemical heating device for point-of-care detection of HIV-1. *PLoS One* 7:e31432. doi: 10.1371/journal.pone.0031432
- Demkin, V. V., Koshechkin, S. I., and Slesarev, A. (2017). A novel real-time PCR assay for highly specific detection and quantification of vaginal *Lactobacilli*. *Mol. Cell Probes* 32, 33–39. doi: 10.1016/j.mcp.2016.11.006
- Han, F., Wang, F., and Ge, B. (2011). Detecting potentially virulent *Vibrio vulnificus* strains in raw oysters by quantitative loop-mediated isothermal amplification. *Appl. Environ. Microbiol.* 77, 2589–2595. doi: 10.1128/AEM.02992-10
- Hoegh, S. V., Skov, M. N., Boye, K., Worning, P., Jensen, T. G., and Kemp, M. (2014). Variations in the *Staphylococcus aureus*-specific *nuc* gene can potentially lead to misidentification of methicillin-susceptible and -resistant *S. aureus*. *J. Med. Microbiol.* 63, 1020–1022. doi: 10.1099/jmm.0.076638-0
- Huehn, S., Eichhorn, C., Urmersbach, S., Breidenbach, J., Bechlers, S., Bier, N., et al. (2014). Pathogenic vibrios in environmental, seafood and clinical sources in Germany. *Int. J. Med. Microbiol.* 304, 843–850. doi: 10.1016/j.ijmm.2014.07.010
- Jarman, S. N. (2004). Amplicon: software for designing PCR primers on aligned DNA sequences. *Bioinformatics* 20, 1644–1645. doi: 10.1093/bioinformatics/bth121
- Kadariya, J., Smith, T. C., and Thapaliya, D. (2014). *Staphylococcus aureus* and staphylococcal food-borne disease: an ongoing challenge in public health. *Biomed. Res. Int.* 2014:827965. doi: 10.1155/2014/827965
- Kalendar, R., Tselykh, T. V., Khassenov, B., and Ramanculov, E. M. (2017). Introduction on using the FastPCR software and the related java web tools for PCR and oligonucleotide assembly and analysis. *Methods Mol. Biol.* 1620, 33–64. doi: 10.1007/978-1-4939-7060-5\_2
- Karmakar, A., Dua, P., and Ghosh, C. (2016). Biochemical and molecular analysis of *Staphylococcus aureus* clinical isolates from hospitalized patients. *Can. J. Infect. Dis. Med. Microbiol.* 2016:9041636. doi: 10.1155/2016/9041636
- Kurosaki, Y., Grolla, A., Fukuma, A., Feldmann, H., and Yasuda, J. (2010). Development and evaluation of a simple assay for Marburg virus detection using a reverse transcription-loop-mediated isothermal amplification method. *J. Clin. Microbiol.* 48, 2330–2336. doi: 10.1128/JCM.01224-09
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

## ACKNOWLEDGMENTS

We thank the High Performance Computing Center (HPCC) at the Shanghai Jiao Tong University for computing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.02860/full#supplementary-material>

- Lim, J. Y., Yoon, J., and Hovde, C. J. (2010). A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J. Microbiol. Biotechnol.* 20, 5–14. doi: 10.4014/jmb.0908.08007
- Liu, J., Bai, L., Li, W., Han, H., Fu, P., Ma, X., et al. (2018). Trends of foodborne diseases in China: lessons from laboratory-based surveillance since 2011. *Front. Med.* 12, 48–57. doi: 10.1007/s11684-017-0608-6
- Mizan, M. F., Jahid, I. K., and Ha, S. D. (2015). Microbial biofilms in seafood: a food-hygiene challenge. *Food Microbiol.* 49, 41–55. doi: 10.1016/j.fm.2015.01.009
- Mori, Y., and Notomi, T. (2009). Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases. *J. Infect. Chemother.* 15, 62–69. doi: 10.1007/s10156-009-0669-9
- Nakyinsige, K., Man, Y. B., and Sazili, A. Q. (2012). Halal authenticity issues in meat and meat products. *Meat Sci.* 91, 207–214. doi: 10.1016/j.meatsci.2012.02.015
- Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., et al. (2000). Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.* 28:E63.
- Okada, K., Chantaroj, S., Taniguchi, T., Suzuki, Y., Roobthaisong, A., Puiprom, O., et al. (2010). A rapid, simple, and sensitive loop-mediated isothermal amplification method to detect toxigenic *Vibrio cholerae* in rectal swab samples. *Diagn. Microbiol. Infect. Dis.* 66, 135–139. doi: 10.1016/j.diagmicrobio.2009.09.004
- O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189
- Parida, M., Sannarangai, S., Dash, P. K., Rao, P. V., and Morita, K. (2008). Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases. *Rev. Med. Virol.* 18, 407–421. doi: 10.1002/rmv.593
- Paudyal, N., Anihouvi, V., Hounhouigan, J., Matsheka, M. I., Sekwati-Monang, B., Amoa-Awua, W., et al. (2017). Prevalence of foodborne pathogens in food from selected African countries - A meta-analysis. *Int. J. Food Microbiol.* 249, 35–43. doi: 10.1016/j.ijfoodmicro.2017.03.002
- Peng, H., Qin, L., Bi, Y., Wang, P., Zou, G., Li, J., et al. (2015). Rapid detection of the common avian leukosis virus subgroups by real-time loop-mediated isothermal amplification. *Virol. J.* 12:195. doi: 10.1186/s12985-015-0430-1
- Ran, G., Ren, L., Han, X., Liu, X., Li, Z., Pang, D., et al. (2016). Development of a rapid method for the visible detection of pork DNA in halal products by loop-mediated isothermal amplification. *Food Anal. Methods* 9, 565–570. doi: 10.1007/s12161-015-0246-z
- Tarr, P. I., Gordon, C. A., and Chandler, W. L. (2005). Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* 365, 1073–1086. doi: 10.1016/s0140-6736(05)71144-2
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Torres, C., Vitalis, E. A., Baker, B. R., Gardner, S. N., Torres, M. W., and Dzenitis, J. M. (2011). LAVA: an open-source approach to designing LAMP



- (loop-mediated isothermal amplification) DNA signatures. *BMC Bioinformatics* 12:240. doi: 10.1186/1471-2105-12-240
- Treven, P. (2015). Strategies to develop strain-specific PCR based assays for probiotics. *Benef. Microb.* 6, 887–898. doi: 10.3920/BM2015.0009
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. doi: 10.1093/nar/gks596
- Wang, X. R., Wu, L. F., Wang, Y., Ma, Y. Y., Chen, F. H., and Ou, H. L. (2015). Rapid detection of *Staphylococcus aureus* by loop-mediated isothermal amplification. *Appl. Biochem. Biotechnol.* 175, 882–891. doi: 10.1007/s12010-014-1328-x
- Wang, Y., Li, D., Li, K., and Ye, C. (2016). Rapid and sensitive detection of *Vibrio parahaemolyticus* and *Vibrio vulnificus* by multiple endonuclease restriction real-time loop-mediated isothermal amplification technique. *Molecules* 21:E1111. doi: 10.3390/molecules21010111
- Xu, W., Zhai, Z., Huang, K., Zhang, N., Yuan, Y., Shang, Y., et al. (2012). A novel universal primer-multiplex-PCR method with sequencing gel electrophoresis analysis. *PLoS One* 7:e22900. doi: 10.1371/journal.pone.0022900
- Yang, L., Fu, S., Peng, X., Li, L., and Song, T. (2014). Identification of pork in meat products using real-time loop-mediated isothermal amplification. *Biotechnol. Biotechnol. Equip.* 28, 882–888. doi: 10.1080/13102818.2014.963789
- Yao, X., Li, P., Xu, J., Zhang, M., Ren, R., Liu, G., et al. (2016). Rapid and sensitive detection of *Didymella bryoniae* by visual loop-mediated isothermal amplification assay. *Front. Microbiol.* 7:1372. doi: 10.3389/fmicb.2016.01372
- Zhao, X., Lin, C. W., Wang, J., and Oh, D. H. (2014). Advances in rapid detection methods for foodborne pathogens. *J. Microbiol. Biotechnol.* 24, 297–312.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Jia, Li, Liu, Lu, Lu, Ma, Li and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.