

# Identification of Multi-repeat Sequences using Genome Mining Approaches for Developing Highly Sensitive Molecular Diagnostic Assay for the Detection of *Chlamydia trachomatis*

Clement Shiluli<sup>1</sup>, Shwetha Kamath<sup>2</sup>, Bernard N. Kanoi<sup>1</sup>, Rachael Kimani<sup>1</sup>, Michael Maina<sup>1</sup>, Harrison Waweru<sup>1</sup>, Moses Kamita<sup>1</sup>, Ibrahim Ndirangu<sup>1</sup>, Hussein M. Abkallo<sup>4</sup>, Bernard Oduor<sup>4</sup>, Nicole Pamme<sup>5</sup>, Joshua Dupaty<sup>3</sup>, Catherine M. Klapperich<sup>3</sup>, Srinivasa Raju Lolabattu<sup>2\*</sup> and Jesse Gitaka<sup>1\*</sup>

<sup>1</sup>Centre for Research in Infectious Diseases, College of Graduate Studies and Research, Mount Kenya University, Thika, Kenya.

<sup>2</sup>Division of Research and Development, Jigsaw Bio Solutions Private Limited, Bangalore, India.

<sup>3</sup>Department of Biomedical Engineering, Boston University, Boston, Massachusetts, USA.

<sup>4</sup>Animal and Human Health Program, International Livestock Research Institute, Nairobi, Kenya.

<sup>5</sup>Department of Materials and Environmental Chemistry, Stockholm University, Sweden

**Keywords:** Identical multi-repeat sequences, *Chlamydia trachomatis*, Isothermal assays

## 1. Summary

*Chlamydia trachomatis* (*C. trachomatis*) is a common sexually transmitted infection (STI). In 2019, the World Health Organization reported about 131 million infections. The majority of infected patients are asymptomatic with cases remaining undetected. It is likely that missed *C. trachomatis* infections contribute to preventable adverse health outcomes in women and children. Consequently, there is an urgent need of developing efficient diagnostic methods. In this study, genome-mining approaches to identify identical multi-repeat sequences (IMRS) distributed throughout the *C. trachomatis* genome were used to design a primer pair that would target regions in the genome. Genomic DNA was 10-fold serially diluted (100pg/ $\mu$ L to  $1 \times 10^{-3}$ pg/ $\mu$ L) and used as DNA template for PCR reactions. The gold standard PCR using 16S rRNA primers was also run as a comparative test, and products were resolved on agarose gel. The novel assay, *C. trachomatis* IMRS-PCR, had an analytical sensitivity of 9.5 fg/ $\mu$ L, representing better sensitivity compared with 16S rRNA PCR (4.31 pg/ $\mu$ L). Our experimental data demonstrate the successful development of lateral flow and isothermal assays for detecting *C. trachomatis* DNA with potential use in field settings. There is a potential to implement this concept in miniaturized, isothermal, microfluidic platforms, and laboratory-on-a-chip diagnostic devices for reliable point-of-care testing.

## 2. Introduction

*Chlamydia trachomatis* (*C. trachomatis*) is one of the most common sexually transmitted infections (STIs) associated with adverse birth and neonatal outcomes such as preterm labor with low birth weight (1). During delivery, the risk of vertical transmission of *C. trachomatis* increases (2). Most patients with *C. trachomatis* infections are asymptomatic, and these cases remain undetected and untreated, further complicating the standard of

### \*Corresponding authors:

Jesse Gitaka

[jgitaka@mku.ac.ke](mailto:jgitaka@mku.ac.ke)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Srinivasa Raju Lolabattu

[raju@jigsawbio.com](mailto:raju@jigsawbio.com)

care in most countries (3). It is, therefore, likely that missed maternal *C. trachomatis* infections contribute to preventable adverse health outcomes among women and children globally, such as endometritis, salpingitis, tubo-ovarian abscesses, pelvic peritonitis, and perihepatitis (4). In 2019 the World Health Organization (WHO) reported about 131 million cases of *C. trachomatis* globally (5).

Traditional methods of *C. trachomatis* testing include cervical cytological examination using Papanicolaou (Pap) smears (6). However, these techniques have many challenges, such as technical difficulties during cultures, labor intensiveness, increased turnaround time, and high cost.

*C. trachomatis* infections are diagnosed by direct and indirect methods (7).

Direct methods detect the presence of *C. trachomatis* in localized infections (7). These methods include culture, antigen tests (Enzyme Immuno Assays (EIA), Direct Fluorescent Antibody (DFA) tests, immune chromatographic tests, Rapid Detection Tests (RDTs), and Nucleic Acid Amplification Tests (NAATs) (7).

Indirect methods are specific to *C. trachomatis* antibodies and may be used for diagnostic evaluation of chronic/invasive infection (Pelvic Inflammatory Disease, lymphogranuloma venereum) and post-infectious complications, like sexually acquired reactive arthritis (SARA) (7).

Direct detection methods, such as detecting *C. trachomatis* using NAATs, have several advantages. They are based on polymerase chain reaction (PCR) and use fluorescence-labeled probes to detect amplification products in real-time, thereby significantly reducing the turn-around time (8). Studies have also shown that *C. trachomatis* PCR using DNA extracted from conjunctival swabs can achieve a detection limit of up to 100 plasmid copies (9).

Other NAATs, such as ligase chain reaction (LCR) and transcription-mediated amplification (TMA), are also sensitive and specific in the detection of *C. trachomatis* in clinical samples (10). Results of these tests can be generated in a few hours when coupled with automated nucleic acid extraction techniques.

However, NAATs such as the Abbot Real-Time CT/NG requires stringent sample transport and storage conditions (11). For example, to achieve accurate and reliable results, samples from asymptomatic women must be stored between 2°C and 30°C and processed within 14 days after collection (11). Symptomatic women's specimens must be thawed, frozen, and stored at the same temperature range (11). These conditions may be challenging, mainly in resource-limited countries where refrigeration facilities are unavailable, especially during specimen collection and transport (12).

When compared to *C. trachomatis* culture, the DFA has a sensitivity of between 95% - 100% (13). However, DFA involves labor-intensive microscopic examination of individual stained specimens, which can be time-consuming when dealing with many samples (13). Also, highly skilled and experienced personnel perform routine microscopic examinations (13). In contrast, EIAs are ideal for testing large numbers of samples and achieve sensitivities of 90% in comparison with culture, but may be less sensitive than DFA and may give false positive tests (13).

Thriving *C. trachomatis* culture depends on isolated vital organisms, and the detection rate in clinical samples ranges between 60%–80%. In addition, culture sensitivity may be affected by inadequate specimen collection, storage, and transport, toxic substances in clinical specimens, and overgrowth of cell cultures by commensal microorganisms (14). *C. trachomatis* culture has an extended turn-around time; it is labor intensive, and different laboratories have various standardization protocols (14).

The use of indirect methods for the detection of *C. trachomatis* is inaccurate in identifying acute infections of the lower genital and anal tract; this is because antibody responses become detectable only after weeks to months and are often less pronounced (14).

Enzyme immunoassays based on the detection of bacterial lipopolysaccharide may cross-react with other gram-negative bacteria leading to false-positive results. The chlamydial antibody response is either delayed or absent in some individuals, which makes most serological tests inadequate (15). In addition, most the *C. trachomatis* infections are asymptomatic and are diagnosed late, resulting in uninterrupted transmission (16). To manage asymptomatic *C. trachomatis* infections, there is an urgent need to develop

efficient diagnostic with adequate sensitivity and specificity, which, when implemented as part of screening programs, can contribute to identifying new cases and controlling the transmission of *C. trachomatis* infections (17). In this study, we developed a novel method based on de novo genome mining strategy that identifies identical multi-repeat sequences (IMRS) in *C. trachomatis* genome for use as both PCR and isothermal amplification assays. The assay has a potential for field deployability due to inherent high sensitivity.

## 3. Materials and Methods

### 3.1 IMRS genome mining algorithm

The primers used in this study were designed based on Identical Multi-Repeat Sequence (IMRS) genome mining algorithm as previously described (18). The *C. trachomatis* genome (NC\_000117.1) was used as the reference to identify identical, repetitive sequence substrings of <30 bases that can be used as forward or reverse primers. NIH's Basic Local Alignment Search Tool (BLAST) was used to evaluate the primer pairs to confirm *C. trachomatis* specificity. The resulting primers were predicted to amplify several fragments of DNA of different sizes derived from different regions of the *C. trachomatis* genome.

### 3.2 *C. trachomatis* Genomic DNA Preparation

*C. trachomatis* DNA was obtained from the American Type Culture Collection (ATCC) (ATCC® VR-885D®, LOT No. 70013611) at a concentration of  $\geq 1 \times 10^6$  copies/ $\mu$ L. The stock DNA was diluted to a concentration of 100 pg/ $\mu$ L ( $8.92 \times 10^4$  copies/ $\mu$ L) and thereafter serially diluted 100 fold and 10 fold in TE buffer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) for the *C. trachomatis* IMRS and 16S rRNA PCR, respectively.

### 3.3 16S rRNA PCR

The 16S rRNA PCR assays was carried out in a reaction mixture containing dNTPs (Thermo Fisher Scientific, Waltham, Massachusetts, USA) (0.2mM), forward and reverse primers (0.01 mM each), Taq Hot-Start DNA polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA) (1.25 U), genomic template DNA, 1  $\mu$ L to a final PCR reaction volume of 25  $\mu$ L. The cycling parameters were as follows: 95°C for 3 min; 40 cycles of: 95°C for 30 s, 56°C for 30 s, 72°C for 30 s; and 72°C for 5 min and a final hold of 4°C.

### 3.4 *C. trachomatis* IMRS PCR

The *C. trachomatis* IMRS assays were carried out in a reaction mixture containing dNTPs (Thermo Fisher Scientific, Massachusetts, USA) (0.2 mM), forward and reverse primers (0.01 mM each), Taq Hot-Start DNA polymerase (Thermo Fisher Scientific, Massachusetts, USA) (1.25 U), genomic template DNA, 1  $\mu$ L to a final PCR reaction of 25  $\mu$ L. The cycling parameters for *C. trachomatis*-IMRS assay was as follows: 95°C for 3 min; 40 cycles of: 95°C for 30 s, 50°C for 30 s, 72°C for 30 s; and 72°C for 5 min and a final hold of 4°C. All PCR products were resolved in 1% agarose gel visualized on a UV Gel illuminator system (Fison Instruments, Glasgow, United Kingdom) under ethidium bromide staining.

### 3.5 Isothermal IMRS Amplification Assay

The Isothermal (Iso) IMRS amplification were performed in a 25  $\mu$ L reaction mixture consisted of *Bst* 2.0 polymerase (640 U/mL) (New England Biolabs, Massachusetts, USA), with 1 $\times$  isothermal amplification buffer, 3.2  $\mu$ M forward primer, and 1.6  $\mu$ M reverse primer (Jigsaw Biosolutions, Bengaluru, India) combined with 10 mM dNTPs (Thermo Fisher Scientific, Massachusetts, USA), 0.4 M Betaine (Sigma-Aldrich, Missouri, USA), molecular-grade water and Ficoll (0.4 g/mL) (Sigma-Aldrich, Missouri, USA). Amplification was carried out at 56°C for 40 min. Amplified products were visualized by gel electrophoresis in 1% gel.

### 3.6 Lower limit of detection

To assess the lower limit of detection (LLOD) of the *C. trachomatis* IMRS PCR assay, genomic DNA was diluted 100-fold from 100 pg/ $\mu$ L ( $8.92 \times 10^4$  copies/ $\mu$ L) to 10<sup>-6</sup> pg/ $\mu$ L

(<1 copies/ $\mu$ L) and 10-fold from 100 pg/ $\mu$ L ( $8.92 \times 10^4$  copies/ $\mu$ L) to  $10^2$  pg/ $\mu$ L (< 1 copies/ $\mu$ L) for the gold standard 16S rRNA PCR. Thereafter, 5 replicates of each dilution were used for the assays. Amplification products were visualized on gel after electrophoresis. To determine the LLOD of the *C. trachomatis* 16S rRNA and IMRS PCR, probit analysis was performed using the ratio of successful reactions to the total number of reactions performed for each assay. Similarly, to assess the LLOD for the *C. trachomatis* Iso IMRS PCR assay, genomic DNA was diluted 10-fold from a starting concentration of  $1.64 \times 10^6$  copies/ $\mu$ L. Thereafter, five replicates of each dilution were used to determine the LLOD by using the ratio of successful reactions to the total number of reactions performed for each assay.

### 3.7 *C. trachomatis*-Lateral Flow Assay

To generate a visual read-out signal of amplicons, the *C. trachomatis*-Lateral Flow Assay (*C. trachomatis*-LFA) was used. Briefly, 2.5  $\mu$ L annealing buffer, dNTPs and NaCl (1.75  $\mu$ L), MgSO<sub>4</sub> (1.2  $\mu$ L), NG 5' biotinylated forward primer (Biotin 5'-"TGCTGCTGCTGATTACGAGCCGA"-3'), *C. trachomatis* Reverse primer, *C. trachomatis* 3' FAM labelled probe (CCACCAATACTCTC/-FAM-3'), *C. trachomatis* Digotexin labelled probe, Ficoll 400 (6.25  $\mu$ L), *C. trachomatis* internal control sequence DNA (2.5  $\mu$ L), ISO Amp III enzyme mix 2.0  $\mu$ L, template DNA 5 $\mu$ L and molecular grade water 1.45  $\mu$ L in a final master mix volume of 25  $\mu$ L was used. Incubation of the LFA strip (Milenia Biotec GmbH, Giessen, Germany) was performed at 65°C for 1 hour and 5  $\mu$ L of reaction mixture was then added to the LFA strip, thereafter, 2 drops of running buffer was also added.

### 3.8 *C. trachomatis* IMRS and 16S rRNA PCR real-time PCR assay

The Quant Studio 5 Real-Time PCR System (with Quant Studio Design and Analysis Desktop Software v1.5.) was used as a reference method (19) for the *C. trachomatis* 16S rRNA PCR and the *C. trachomatis* IMRS-PCR and for determining the sensitivity of the *C. trachomatis* IMRS and 16S rRNA PCR primers for detecting *C. trachomatis* DNA. The genomic DNA was serially diluted 10-fold starting concentration of  $10^6$  genome copies/ $\mu$ L. The final real-time PCR master mix volume was 10  $\mu$ L in triplicate and consisted of the following components: 5  $\mu$ L SYBR Green qPCR Master Mix (Thermo fisher, Massachusetts, USA), 1  $\mu$ L forward and reverse IMRS primer mix, 2.5  $\mu$ L template genomic DNA and 1.5  $\mu$ L molecular grade water. The amplification cycling conditions were 50°C for 2 min; 95°C for 10 min; 40 cycles of 95°C for 15 s and 60°C for 30 s.

### 3.9 Cloning and characterization of *C. trachomatis*-IMRS amplicons

Gene cloning was performed to confirm the sequences of amplicons obtained from the *C. trachomatis* IMRS PCR assay. *Chlamydia trachomatis* gDNA was amplified using Assembly\_IMRS-F (TTCCGGATGGCTCGAGTTTTTCAGCAAGATTGCCTGCCTGCTGATTACGAGCCG **A**) and Assembly IMRS-R (AGAATATTGTAGGAGATCTTCTAGAAAGATT**GTAGGAGGAGCCTCTTAGAGA** **A**) primers. The underlined bold sequences correspond to the IMRS primers for amplifying the *C. trachomatis* genome whereas the bold underline in the primers corresponds to sequences in the cloning vector. The resulting amplicons was resolved on 2% agarose to confirm the fragment size and subsequently purified using the PureLink™ PCR purification kit (ThermoFisher). The purified amplicon was then ligated into pJET1.2/ blunt vector (ThermoFisher) using the NEBuilder® HiFi DNA Assembly kit (NEB) as per the manufacturer's instruction. The resulting NEBuilder HiFi DNA Assembly product was transformed into NEB 5-alpha Competent *E. coli* (NEB #C2987, NEB) following the manufacturer's instructions. Transformed colonies were randomly selected, DNA extracted and Sanger-sequenced using the universal pJET1.2 forward sequencing primer (CGACTCACTATAGGGAGAGCGGC) and pJET1.2 forward



sequencing primer (AAATACATCCATTCATGAA). The resulting nucleotides were trimmed and analysed using SnapGene software (GSL Biotech; available at [snapgene.com](http://snapgene.com)), aligned to check for similarity or “clonal” differences and BLAST used to check for similarity with the *C. trachomatis* genome.

### 3.10 Clinical samples

Vaginal swab samples collected from a cohort of women aged between 19 – 49 years enrolled in an STI screening study at the Kenyatta National Hospital in Nairobi County in 2022 were used in this study. The *C. trachomatis*-16S rRNA conventional PCR assay was used to confirm *C. trachomatis* infections. Thereafter, *C. trachomatis* positive samples were used to validate the *C. trachomatis*-IMRS primers. Participants were notified of results directly and confidentially by study staff, and were treated for *C. trachomatis* infection. Use of these samples and study proposal was approved by the Mount Kenya University Ethical Review Committee (MKU/ERC/1649).

### 3.11 Data analysis

Graphs were plotted with GraphPad Prism version 7.0 (GraphPad Software, San Diego, CA). The mean, and SD values were calculated with Excel 2016. To determine the LLOD of *C. trachomatis*-IMRS and *C. trachomatis*-16S rRNA PCR assays (the concentration at which genomic *Chlamydia trachomatis* DNA is detected with 95% confidence), probit regression analyses were performed in Excel 2016. Genomic *Trichomonas vaginalis* and *Treponema pallidum* DNA were used to determine the specificity of the *C. trachomatis*-IMRS primers for other STIs. Statistical analyses were performed using *t*-test of GraphPad Prism version 7.0. for two-tailed distribution.  $P < 0.05$  was considered as significant.

## 4. Results

### 4.1 Design and Distribution of IMRS Primer Sets on *Chlamydia trachomatis* Genome

A total of 6 repeats (Table 1) were identified using the IMRS-based genome mining algorithm [10], which could be used as forward and reverse primers for an amplification assay. These selected primers (F 5'- TGCTGCTGCTGATTACGAGCCGA -3' and R 5'-TGTAGGAGGAGCCTCTTAGAGAA - 3'), as were depicted using a circos plot (Fig 1), were found to be present at various loci of the sense and antisense strands, allowing them to serve interchangeably as forward or reverse primers.

### 4.2 Amplification of Sequences on *Chlamydia trachomatis* Genome using Identical Multi-Repeat Primers and Gold Standard 16S rRNA PCR

To confirm the ability and specificity of the IMRS primers to amplify the targeted regions on the *C. trachomatis* genome, serially diluted DNA was used as a template for PCR amplification. The IMRS primers could detect *C. trachomatis* genomic DNA from a concentration of as low as 0.1 fg/ $\mu$ L (Fig 1B), whereas the gold standard 16S rRNA primers (Fig 1C) could only detect *C. trachomatis* genomic DNA down to a concentration of 10 fg/ $\mu$ L. This demonstrates that the IMRS-based PCR assay has a 100-fold higher sensitivity compared to the gold standard 16S rRNA PCR assay.

Real-time PCR was also performed using serially diluted genomic *C. trachomatis* DNA as a template and *C. trachomatis*-IMRS primers (Table 2A), as well as conventional *C. trachomatis*-16S rRNA primers (Table 2B). The mean Ct values at each respective dilution were used to plot amplification curves (Fig 2A for *C. trachomatis*-IMRS primers and Fig 2B for *C. trachomatis*-16S rRNA primers).

### 4.3 Determination of the Lower Limit of Detection (LLOD) for IMRS PCR and Gold Standard 16S rRNA PCR assay

To determine the lowest limit of detection (LLOD) of the IMRS PCR relative to the gold standard 16S rRNA PCR, probit statistic was performed using *C. trachomatis* genomic DNA serially diluted 100 fold (Table 3A) and 10-fold (Table 3B) and used as template for the *C. trachomatis*-IMRS and 16S RNA PCR. Fig 3A shows the probit plot for the *C. trachomatis*-IMRS PCR assay, and Fig 3B shows the probit plot for the gold standard 16S

rRNA PCR assay. The LLOD was calculated as the concentration at which *C. trachomatis* DNA can be detected with 95% confidence. Probit analysis estimation for *C. trachomatis*-IMRS PCR, Coefficient  $\chi = -3.6494$ ,  $P$ -value 0.7043 (Table 3C). As indicated, the IMRS primers for *C. trachomatis* had an LLOD = 9.5 fg/ $\mu$ L, Fig 3A. Probit analysis estimation for 16S rRNA PCR  $\chi = -7.2101$ ,  $P$ -value 0.9978, Table 3C. As indicated, gold standard primers for *C. trachomatis* had an LLOD = 4.31 pg/ $\mu$ L. Table 3C shows the statistics obtained from the Probit analysis indicating that the *C. trachomatis*-IMRS PCR assay had increased sensitivity compared to the gold standard 16S rRNA PCR assay.

#### 4.4 Isothermal Amplification of Genomic *Chlamydia trachomatis* DNA

Serially diluted genomic DNA was used to perform Iso-*C. trachomatis*-IMRS isothermal amplification. As shown in Fig 4A, the reaction products were visualized on a 1% gel. The Iso-*C. trachomatis*-IMRS assay successfully amplified *C. trachomatis* DNA down to  $1.64 \times 10^6$  genome copies/ $\mu$ L using isothermal amplification. The LLOD for the *C. trachomatis*-Iso-IMRS assay was estimated at 0.3162 ng/ $\mu$ L (Fig 4C).

#### 4.5 *Chlamydia trachomatis* Lateral Flow Assay (LFA)

A visual readout of diluted genomic DNA was developed using modified primers and labelled probes, as shown in Fig 4B. The LFA readout of the amplification products was successful indicating the potential for the Iso-IMRS-based assay to be used in field settings.

#### 4.6 Plasmid *C. trachomatis*-DNA concentration in ng/ $\mu$ L of transformed *E. coli* cells

To validate the exact regions that were amplified by the *C. trachomatis*-IMRS primers, we cloned the amplicon into blunt cloning vectors and transformed into electrocompetent *E. coli* cells and plated onto agar plates. DNA from eight (Figure 5) transformed *E. coli* cells was extracted and sequenced. Multiple sequencing alignment confirmed *C. trachomatis* sequences. These results suggested that the *C. trachomatis*-IMRS primers were specific for targets within the genome.

#### 4.7 Validation of the *C. trachomatis*-IMRS primers using clinical samples

A total of 16 *Chlamydia trachomatis* positive DNA samples from a cross-sectional study at the Kenyatta National Hospital were used to evaluate the reliability of the *C. trachomatis*-IMRS primers for detecting *C. trachomatis* DNA using RT-PCR assay as shown in Fig 6. The demographic information for recruited participants has been described in Table 4. The *C. trachomatis*-IMRS primers showed excellent concordance with the results obtained from conventional PCR diagnosis, successfully identifying *C. trachomatis* infections in infected samples.

#### 4.8 Specificity and sensitivity of the *C. trachomatis*-IMRS primers

As shown in Fig 7, the *C. trachomatis*-IMRS primers were specific only for *C. trachomatis* DNA after amplification of *Treponema pallidum* (TP) and *Trichomonas vaginalis* (TV) genomic DNA using *C. trachomatis*-IMRS primers. *C. trachomatis* genomic DNA was used as a positive control. Also, compared to the gold standard conventional 16S-rRNA PCR, the *C. trachomatis*-IMRS PCR reliably detected genomic DNA at a concentration of 1 fg/ $\mu$ L (Fig 1B).

## 5. Discussion

Traditional techniques for identifying *C. trachomatis* infections have low sensitivity, require the collection of invasive specimens, require long duration for performance and reporting, and are expensive (17). As a consequence, there is a need to develop novel tests that are sensitive and specific. In the present study, we demonstrate the use of a deep genome mining strategy to identify identical multiple repeat sequences that could be used as robust primers for a novel nucleic acid-based test that is a highly sensitive test against *C. trachomatis*. Specifically, IMRS forward and reverse primers initiate amplification at multiple loci across the genome. This, in turn, improves the overall analytical sensitivity

by generating more amplicons (21-22) was observed that, compared to the gold standard 16S rRNA PCR, amplification of specific sequences on the *C. trachomatis* genome using IMRS primers was more sensitive, generating a large number of amplicons of varying sizes hence a lower limit of detection of 0.0095 pg/mL (8.4 genome copies/mL). Our results are comparable to the Chlamydial Roche Amplicor Real-Time Quantitative PCR with a detection limit of 200 genome copies/ml (23). However, this assay targets up to 10 copies of chlamydial plasmid.

We further showed that isothermal amplification using IMRS PCR primers is reliable and sensitive for detecting *C. trachomatis*. The *C. trachomatis*-Iso-IMRS assay detected DNA up to 8.9 genome copies per mL. The Iso-IMRS assay was more sensitive than a Loop-mediated isothermal assay based on the *ompA* and *orf1* genes which reported a limit of detection of 50 copies per mL (24).

Our study successfully developed and tested a novel Lateral Flow Assay testing platform to identify *C. trachomatis* to a concentration of 10 pg/mL (8.8 genome copies/ml) (Fig 6). Our finding was different from a study that developed a visual read-out of *C. trachomatis*-LAMP results, a gold nanoparticle-based lateral flow biosensor that reported a detection limit of 50 copies/ml after an incubation of 45 minutes (20).

Compared to nucleic acid methods that are used for the detection of *C. trachomatis* that are mostly suboptimal (22), the *C. trachomatis*-IMRS PCR primers were specific and sensitive when used for the identification of *C. trachomatis* DNA.

## 6. Conclusion

Put together; here we show that the IMRS algorithm can serve as a platform technology for designing primers that are sensitive and specific for *C. trachomatis*. This platform has potential application in other bacterial and non-bacterial pathogens and could significantly improve future disease diagnostics procedures.

### Ethical Statement

Use of these samples and study proposal was approved by the Mount Kenya University Ethical Review Committee (MKU/ERC/1649).

### Data Accessibility

All data generated or analysed during this study are included in this published article [a supplementary information files].

DNA sequences: Genbank accessions AE001273 AE001275 AEC  
<https://www.ncbi.nlm.nih.gov/nucleotide/AE001273.1/>

### Competing Interests

We have no competing interests.

### Funding Statement

This research was supported by the Royal Society, Future Leaders African Independent Researchers (FLAIR) Scheme (FLR\R1\201314) to JG

### Authors' Contributions

C.S.: conceptualization, data curation, formal analysis, methodology, writing—original draft; S.K.: data curation, formal analysis, investigation, methodology, supervision, visualization, writing—original draft, writing—review and editing; B.N.K.: investigation, writing—review and editing; R.K.: resources, writing—review and editing; M.M.: resources, writing—review and editing; H.W.: resources, writing—review and editing; M.K.: resources, writing—review and editing; I.N.: resources, writing—review and editing; H.M.A.: resources, writing—review and editing; B.O.: funding acquisition, writing—review and editing; N.P.: resources, writing—review and editing; J.D.: writing—review and editing; C.M.K.: resources, writing—review and editing; S.R.L.: methodology, writing—review and editing; J.G.: conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, writing—original draft, writing—review and editing.

### Acknowledgement

We thank Dr. Karen Muthembwa and all clinicians working at the STI Clinic in Kenyatta National Hospital

## References

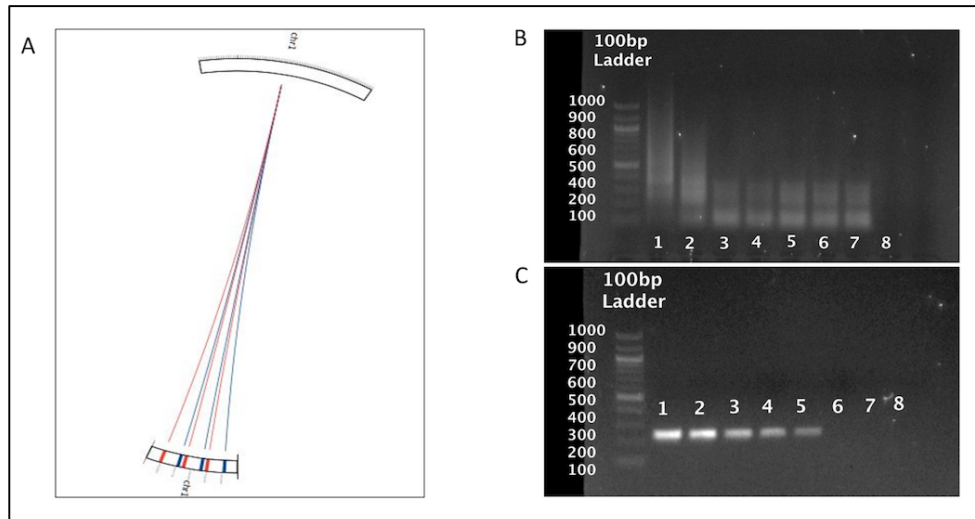
1. Wynn A, Mussa A, Ryan R, Hansman E, Simon S, Bame B, et al. 2022 Evaluating the diagnosis and treatment of Chlamydia trachomatis and Neisseria gonorrhoeae in pregnant women to prevent adverse neonatal consequences in Gaborone, Botswana: protocol for the Maduo study. *BMC Infect Dis.* **1;22(1)**. DOI: [10.1186/s12879-022-07093-z](https://doi.org/10.1186/s12879-022-07093-z)
2. Gadoth A, Shannon CL, Hoff NA, Mvumbi G, Musene K, Okitolonda-Wemakoy E, et al. 2020 Prenatal chlamydial, gonococcal, and trichomonal screening in the Democratic Republic of Congo for case detection and management. *Int J STD AIDS* [Internet]. [cited 2023 Jul 11];**31(3)**:221–9. Available from: /pmc/articles/PMC7534514/ DOI: [10.1177/0956462419888315](https://doi.org/10.1177/0956462419888315)
3. Armstrong-Mensah EA, Ebiringa D-P, Whitfield K, Coldiron J. 2021 Genital Chlamydia Trachomatis Infection: Prevalence, Risk Factors and Adverse Pregnancy and Birth Outcomes in Children and Women in sub-Saharan Africa. *Int J Matern Child Heal AIDS* [Internet]. [cited 2023 Jul 11];**10(2)**:251–7. Available from: <https://mchandaids.org/index.php/IJMA/article/view/523> DOI: [10.21106/ijma.523](https://doi.org/10.21106/ijma.523)
4. Olaleye AO, Babah OA, Osuagwu CS, Ogunsola FT, Afolabi BB. 2020 Sexually transmitted infections in pregnancy – An update on Chlamydia trachomatis and Neisseria gonorrhoeae. *Eur J Obstet Gynecol Reprod Biol.* 2020 **1;255**:1–12. DOI: [10.1016/j.ejogrb.2020.10.002](https://doi.org/10.1016/j.ejogrb.2020.10.002)
5. World Health Organization. Sexually transmitted infections.2019 Vol. **27**. p. 9–10. <https://apps.who.int/iris/bitstream/handle/10665/324797/WHO-CDS-HIV-19.7-eng.pdf>
6. Budai I. 2007 Chlamydia trachomatis: Milestones in clinical and microbiological diagnostics in the last hundred years. *Acta Microbiologica et Immunologica Hungarica.* Vol. **54**,p. 5–22. DOI: [10.1556/AMicr.54.2007.1.2](https://doi.org/10.1556/AMicr.54.2007.1.2)
7. Meyer T.2016 Diagnostic procedures to detect chlamydia trachomatis infections [Internet]. *Microorganisms. Multidisciplinary Digital Publishing Institute (MDPI)*;[cited 2023 Jul 11]. Vol. **4**, Available from: /pmc/articles/PMC5039585/ DOI: [10.3390/microorganisms4030025](https://doi.org/10.3390/microorganisms4030025)
8. Böhm I, Gröning A, Sommer B, Müller HW, Krawczak M, Glaubitz R. A German 2009 Chlamydia trachomatis screening program employing semi-automated real-time PCR: results and perspectives. *J Clin Virol* [Internet]. [cited 2023 Jul 11];**46** Suppl 3(SUPPL. 3). Available from: <https://pubmed.ncbi.nlm.nih.gov/20129071/> DOI: [10.1016/S1386-6532\(09\)70298-7](https://doi.org/10.1016/S1386-6532(09)70298-7)
9. Elnifro EM, Storey CC, Morris DJ, Tullo AB. 1997 Polymerase chain reaction for detection of Chlamydia trachomatis in conjunctival swabs. *Br J Ophthalmol* [Internet]. [cited 2023 Jul 11];**81(6)**:497–500. Available from: <https://pubmed.ncbi.nlm.nih.gov/9274416/> DOI: [10.1136/bjo.81.6.497](https://doi.org/10.1136/bjo.81.6.497)
10. Watson EJ, Templeton A, Russell I, Paavonen J, Mardh PA, Sary A, et al. 2002 The accuracy and efficacy of screening tests for Chlamydia trachomatis: A systematic review., *Journal of Medical Microbiology.* Lippincott Williams and Wilkins; Vol. **51**. p. 1021–31. DOI: [10.1099/0022-1317-51-12-1021](https://doi.org/10.1099/0022-1317-51-12-1021)
11. Masek BJ, Arora N, Quinn N, Aumakhan B, Holden J, Hardick A, et al. 2009 Performance of three nucleic acid amplification tests for detection of Chlamydia trachomatis and Neisseria gonorrhoeae by use of self-collected vaginal swabs obtained via an internet-based screening program. *J Clin Microbiol* [Internet]. [cited 2023 Jul 11];**47(6)**:1663–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/19386838/> DOI: [10.1128/JCM.02387-08](https://doi.org/10.1128/JCM.02387-08)
12. Papp JR, Schachter J, Gaydos CA, Van Der Pol B. 2014 Recommendations for the laboratory-based detection of Chlamydia trachomatis and Neisseria gonorrhoeae - 2014. *MMWR Recomm Reports* [Internet]. Mar 3 [cited 2023 Jul 11];**63(1)**:1. Available from: /pmc/articles/PMC4047970/
13. Lanjouw E, Ouburg S, de Vries HJ, Sary A, Radcliffe K, Unemo M. 2015 European guideline on the management of Chlamydia trachomatis infections. *Int J STD AIDS.* 2016 Apr **1;27(5)**:333–48.
14. CDC. MMWR Chlamydia testing. Guidelines [Internet]. 2002;**51**:48. Available from: [papers3://publication/uuid/3396A926-EB58-4F42-887F-E494892D6FEB](https://pubmed.ncbi.nlm.nih.gov/12111111/)



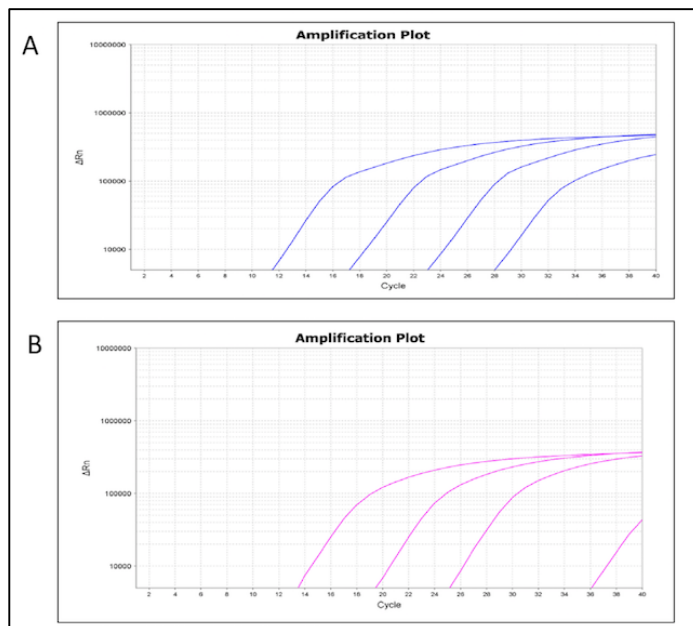
15. de Haro-Cruz MJ, Guadarrama M, Acevedo S, Lopez-Hurtado M, Escobedo-Guerra MR, Guerra-Infante FM. 2019 Obtaining an ELISA test based on a recombinant protein of *Chlamydia trachomatis*. *Int Microbiol*. **22(4)**:471–8. DOI: [10.1007/s10123-019-00074-4](https://doi.org/10.1007/s10123-019-00074-4)
16. Bianchi S, Frati ER, Canuti M, Colzani D, Fasoli E, Amendola A, et al. 2016 Molecular epidemiology and genotyping of *Chlamydia trachomatis* infection in a cohort of young asymptomatic sexually active women (18–25 years) in Milan, Italy. *J Prev Med Hyg* [Internet]. [cited 2023 Jul 11];**57(3)**:E128–34. Available from: [/pmc/articles/PMC5139607/](https://pmc/articles/PMC5139607/)
17. Shetty S, Kouskouti C, Schoen U, Evangelatos N, Vishwanath S, Satyamoorthy K, et al. 2021 Diagnosis of *Chlamydia trachomatis* genital infections in the era of genomic medicine [Internet]. *Brazilian Journal of Microbiology*. Brazilian Society of Microbiology; 2021 [cited 2023 Jul 11]. **Vol. 52**, p. 1327–39. Available from: [/pmc/articles/PMC8221097/](https://pmc/articles/PMC8221097/) DOI: [10.1007/s42770-021-00533-z](https://doi.org/10.1007/s42770-021-00533-z)
18. Raju LS, Kamath S, Shetty MC, Satpathi S, Mohanty AK, Ghosh SK, et al. 2019 Genome Mining–Based Identification of Identical Multirepeat Sequences in *Plasmodium falciparum* Genome for Highly Sensitive Real-Time Quantitative PCR Assay and Its Application in Malaria Diagnosis. *J Mol Diagnostics*. ;**21(5)**:824–38. DOI: [10.1016/j.jmoldx.2019.04.004](https://doi.org/10.1016/j.jmoldx.2019.04.004)
19. Graham EK, Drobac J, Thompson J, Loten M, Gopalakrishnan A. 2018 Evaluation of the PowerQuant<sup>®</sup> System on the QuantStudio<sup>™</sup> 5 Real-Time PCR System. *White Pap Promega Corp.*;1–12.
20. Chen X, Zhou Q, Tan Y, Wang R, Wu X, Liu J, et al. 2022 Nanoparticle-Based Lateral Flow Biosensor Integrated With Loop-Mediated Isothermal Amplification for Rapid and Visual Identification of *Chlamydia trachomatis* for Point-of-Care Use. *Front Microbiol*; **13**(July):1–13. DOI: [10.3389/fmicb.2022.914620](https://doi.org/10.3389/fmicb.2022.914620)
21. Balows A. Manual of clinical microbiology 8th edition. 2003 *Diagn Microbiol Infect Dis*. **47(4)**:625–6.
22. Kulkarni S, Bala M, Risbud A. 2015 Performance of tests for identification of *Neisseria gonorrhoeae*., *Indian Journal of Medical Research*. Vol. **142**. p. 833–5. DOI: [10.4103/0971-5916.160721](https://doi.org/10.4103/0971-5916.160721)
23. Wiggins R, Graf S, Low N, Horner PJ. 2009 Real-time quantitative PCR to determine chlamydial load in men and women in a community setting. *J Clin Microbiol* [Internet]. [cited 2023 Jul 11];**47(6)**:1824–9. Available from: <https://journals.asm.org/doi/10.1128/jcm.00005-09>
24. Chen X, Zhou Q, Yuan W, Shi Y, Dong S, Luo X. 2023 Visual and rapid identification of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* using multiplex loop-mediated isothermal amplification and a gold nanoparticle-based lateral flow biosensor. *Front Cell Infect Microbiol* [Internet]. [cited 2023 Jul 11];**13**:**1067554**. Available from: [/pmc/articles/PMC10011439/](https://pmc/articles/PMC10011439/) DOI: [10.3389/fcimb.2023.1067554](https://doi.org/10.3389/fcimb.2023.1067554)

## LIST OF FIGURES

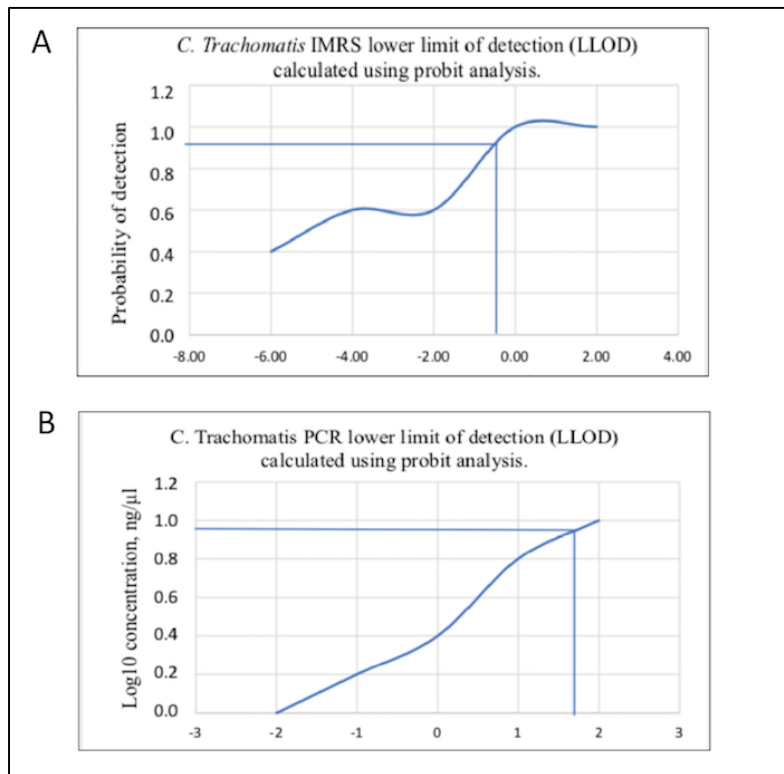
### FIGURES



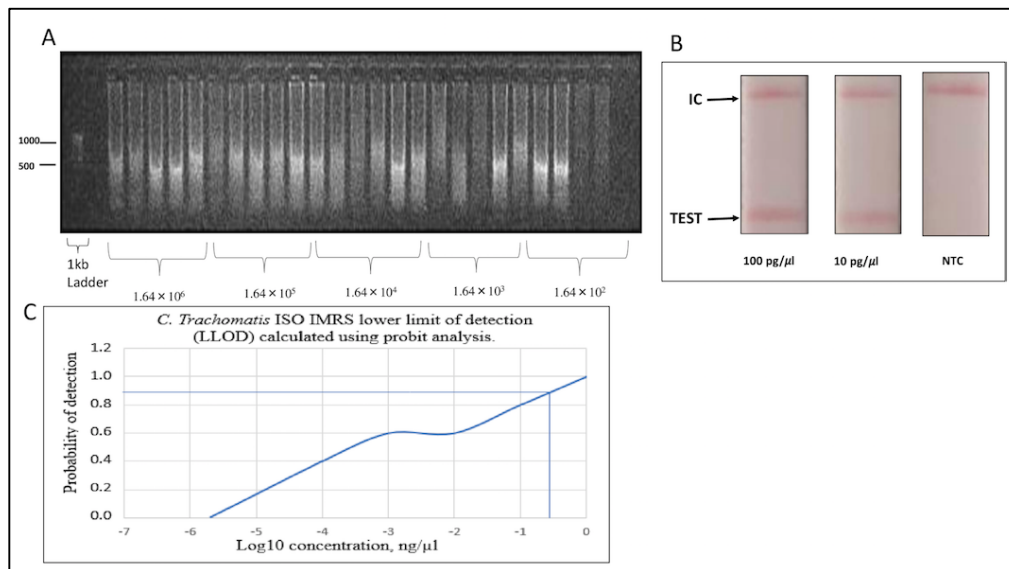
**Fig 1** *C. trachomatis*-IMRS primer targets and gel images from the *C. trachomatis* IMRS and conventional 16S rRNA PCR assay: (A) Circos plot for the distribution of identical multi repeat sequence (IMRS) primers in the *Chlamydia trachomatis* genome. *Chlamydia trachomatis* IMRS primer A (blue lines) IMRS primer B (red lines) both have 6 repeats. Image of 10-fold serially diluted (1 – 100, 2 – 10, 3 – 1, 4 – 0.1, 5 – 0.01, 6 – 0.01, 7 – 0.001 and 8 - NTC) (pg/ $\mu$ l) genomic *Chlamydia trachomatis* DNA amplicons resolved on 1% gel using IMRS primers (B) and gold standard 16S rRNA PCR (C).



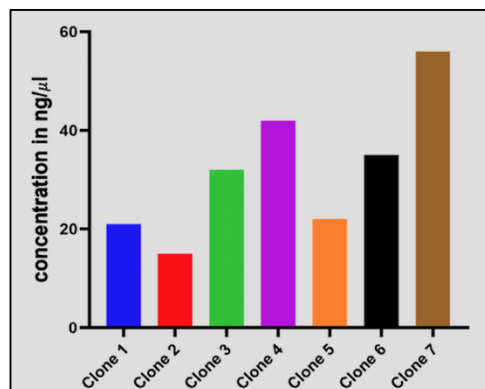
**Fig 2:** Real-time PCR amplification plots from the *C. trachomatis*- IMRS (A) and the 16S rRNA (B) assays using serially diluted *Chlamydia trachomatis* genomic DNA.



**Fig 3:** Probit regression analysis to estimate the lower limit of detection for the *C. trachomatis*-IMRS primers and the 16S rRNA PCR. Probit analysis estimation for *C. trachomatis*-IMRS (A). As indicated, the IMRS primers for *C. trachomatis* had an LLOD = 9.5 fg/μl. B: Probit analysis estimation for 16S rRNA PCR. As indicated, gold standard primers for *C. trachomatis* had an LLOD = 4.31 pg/μl.

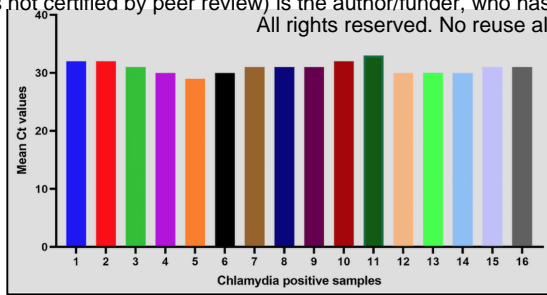


**Fig 4:** *C. trachomatis*-Iso-IMRS, lateral flow assay and estimation of the lower limit of detection of the isothermal assay. (A) Gel image of *C. trachomatis*-Iso IMRS assay products visualized on 1% gel. 5 replicates of each dilution served as DNA template for the *C. trachomatis*-Iso IMRS. DNA concentration is in genome copies per  $\mu\text{l}$ . (B) Visual read-out detection of serially diluted *Chlamydia trachomatis* DNA using the lateral flow assay. Amplicons were incubated at  $65^\circ\text{C}$  for 1 hour and transferred onto strips as indicated. IC – Internal Control, NTC – Non Template Control. (C) Probit analysis estimation for *C. trachomatis* Iso-IMRS. As indicated, the IMRS primers for *C. trachomatis* had an LLOD = 0.3162  $\text{ng}/\mu\text{l}$ .

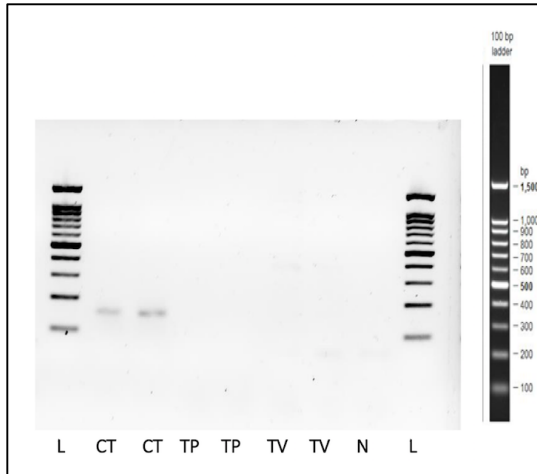


**FIG 5:** Transformed *E.coli* cells expressing *Chlamydia trachomatis* sequences amplified using *C. trachomatis*-IMRS primers.





**FIG 6:** Mean Ct values of PCR confirmed 16 clinical DNA samples from RT-PCR assay that were used to validate the *C. trachomatis*-IMRS PCR primers.



**Fig 7:** Gel image of PCR products obtained after amplification of *Treponema pallidum* (TP) and *Trichomonas vaginalis* (TV) genomic DNA using *C. trachomatis*-IMRS primers. *C. trachomatis* genomic DNA was used as a positive control. N = Negative control, L = Ladder.

## LIST OF TABLES

**TABLE 1:** IMRS Primer target regions on the *Chlamydia trachomatis* bacterial genome

No.	Amplification regions	Expected sequence fragment
1	531375 - 531511	137bp
2	531375 - 531661	287bp
3	531375 - 531814	440bp
4	531525 - 531661	137bp
5	531525 - 531814	290bp
6	531675 - 531814	140bp

**Table 2:** Genomic DNA dilution to determine the sensitivity of the *C. trachomatis*-IMRS primers using Real time PCR. Serially diluted *Chlamydia trachomatis* genomic DNA served as amplification templates for the *C. trachomatis*-IMRS (A) and 16S rRNA primers (B).

**A**

Concn of DNA (genome copies/ $\mu$ l)	Ct value	STD.DEV
$1 \times 10^4$	12.298	0.541
$1 \times 10^3$	17.674	1.244
$1 \times 10^2$	23.639	0.349
$1 \times 10^1$	28.265	0.275

**B**

Concn of DNA (genome copies/ $\mu$ l)	Ct value	STD.DEV
$1 \times 10^4$	14.64	0.902
$1 \times 10^3$	20.584	1.599
$1 \times 10^2$	26.607	1.056
$1 \times 10^1$	36.527	0.249

Table 3: *Chlamydia trachomatis* genomic DNA was serially diluted 100 folds (A) and 10-folds (B) and used as template for the *C. trachomatis*-IMRS and 16S rRNA PCR to estimate the lower limit of detection. Table (C) shows the statistics obtained from the Probit analysis.

**A**

Serial dilutions (pg/μl)	Replicates (5)
100	5/5
1	5/5
0.01	4/5
0.0001	5/5
0.000001	0/5
0.00000001	0/5

**B**

Serial dilutions (pg/μl)	Replicates (5)
100	5/5
10	4/5
1	3/5
0.1	1/5
0.01	1/5
0.001	0/5

**C**

Assay	$\chi$ Coefficient	P-Value
16S rRNA	-7.2101	0.9978
IMRS	-3.6494	0.7043



**Table 4: Demographic data for participants**

<b>Variable</b>	<b>Chlamydia positive, n = 16,(%)</b>	<b>Chlamydia negative, n = 187, (%)</b>	<b>P-Value</b>
<i>Age Category</i>			
10 - 19	0, 0	2, 1	0.9467
20 - 29	6, 38	63, 34	
30 - 39	9, 56	105, 56	
40 - 49	1, 6	17, 9	
<i>Marital status</i>			
Married	13, 81	161, 86	0.5949
Not Married	3, 19	26, 14	
<i>Level of Education</i>			
Primary	1, 6	14, 7	0.5013
Secondary	8, 50	66, 35	
Tertiary	7, 44	107, 57	
<i>Employment status</i>			
Employed	8, 50	55, 29	0.0875
Not employed	8, 50	132, 71	
<i>Parity</i>			
Primipara	1, 6	39, 21	0.3702
Multipara	14, 88	138, 74	
Grandmultipara	1, 6	10, 5	
<i>Gestational Age</i>			
First trimester	0, 0	11, 6	0.0637
Second trimester	7, 44	37, 20	
Third trimester	9, 56	139, 74	
<i>HIV Status</i>			
Positive	1, 6	1, 1	0.0263
Negative	15, 94	186, 99	
<i>Miscarriage</i>			
None	9, 56	111, 59	0.8421
Once	5	52, 28	
Twice	2	14, 7	
Thrice	0, 0	7, 11	
Quadruple	0, 0	3, 2	