(12) STANDARD PATENT (11) Application No. **AU 2007266857 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Mycoplasma genitalium detection assay based on the MG219 gene

(51) International Patent Classification(s)
C12Q 1/68 (2006.01) ***C12Q 1/70*** (2006.01)

(21) Application No: **2007266857** (22) Date of Filing: **2007.05.24**

(87) WIPO No: **WO07/138271**

(30) Priority Data

(31) Number	(32) Date	(33) Country
0610522.5	**2006.05.26**	**GB**

(43) Publication Date: **2007.12.06**
(44) Accepted Journal Date: **2013.01.10**

(71) Applicant(s)
Health Protection Agency

(72) Inventor(s)
Chalker, Vicki;Ison, Cathy

(74) Agent / Attorney
Freehills Patent Attorneys, Level 43 101 Collins Street, Melbourne, VIC, 3000

(56) Related Art
MUSATOVOVA, O. et al. FEMS Microbiology Letters 2003, vol. 229, pages 73-81

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 December 2007 (06.12.2007)

PCT

(10) International Publication Number
WO 2007/138271 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C12Q 1/70* (2006.01)

(21) International Application Number:
PCT/GB2007/001913

(22) International Filing Date: 24 May 2007 (24.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0610522.5 26 May 2006 (26.05.2006) GB

(71) Applicant *(for all designated States except US)*: **HEALTH PROTECTION AGENCY** [GB/GB]; 61 Colindale Avenue, London NW9 5DF (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **CHALKER, Vicki** [GB/GB]; Health Protection Agency, 61 Colindale Avenue, London NW9 5DF (GB). **ISON, Cathy** [GB/GB]; Health Protection Agency, 61 Colindale Avenue, London NW9 5DF (GB).

(74) Agents: **MACLEAN, Martin, Robert** et al.; Mathys & Squire, 120 Holborn, London EC1N 2SQ (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** MYCOPLASMA GENITALIUM DETECTION ASSAY BASED ON THE MG219 GENE

(57) **Abstract:** There is provided a method for detecting *M. genitalium* nucleic acid in a sample, comprising: (i) amplifying a nucleic acid sequence comprising a fragment of SEQ ID NO: 1 (Mg219 gene); and (ii) detecting said amplified nucleic acid sequence.

WO 2007/138271 A1

MYCOPLASMA GENITALIUM DETECTION ASSAY BASED ON THE MG219 GENE

The present invention relates to detection of *Mycoplasma sp.*, in particular, *Mycoplasma genitalium*, and to reagents and kits therefor.

The Class Mollicutes contains Gram-positive bacteria that lack a cell wall, including the genera *Mycoplasma, Acholeplasma* and *Ureaplasma.* There are over 100 species of *Mycoplasma* alone, and 17 species of Mollicutes are considered a part of the human flora – including *M. pneumoniae, M. genitalium, M. fermentans* and *M. penetrans.*

Human and mammalian *Mycoplasmas* infect cells comprising mucus membranes and can have strict host and tissue specificities. By way of example, the primary tissue infected by *M. pneumoniae* is the respiratory tract, whereas *M. genitalium* primarily infects the urogenital tract.

M. genitalium is sexually transmissible, with a transmission rate that is believed to be similar to *Chlamydia trachomatis.* Indeed, clinicians working in the field have described *M. genitalium* as "the next *Chlamydia*". *M. genitalium* is a causative agent in urogenital tract diseases, in particular non-gonococcal urethritis (NGU). *M. genitalium* has also been implicated in pelvic inflammatory disease (PID) and infertility in women. Extragenital infections with *M. genitalium* may also occur, for example in the respiratory tract, in the eye, and in synovial fluid – leading to sexually acquired reactive arthritis (SARA).

Antibiotics used for the treatment of mycoplasmal infections belong to tetracyclines, macrolides-lincosamides and fluoroquinolones. These products are highly active *in vitro* against *Mycoplasmas.* However, some of these antibiotics have a differential activity according to species, and acquired resistance has been reported, mainly in genital *Mycoplasmas.* By way of example, *M. genitalium* is resistant to chloramphenicol. Most mycoplasmal infections are treated using adapted antibiotics, but they may be difficult to treat in immunosuppressed patients.

Mycoplasmas such as *M. genitalium* are very difficult to study by classical genetic methods, both because of their fastidious growth requirements and as a consequence of the absence of selectable markers. As a consequence, isolates have been very difficult to obtain.

Furthermore, the antigenic relationship between *M. pneumoniae* and *M. genitalium* leads to cross-reactions, which significantly hamper the use of serology for diagnostic and epidemiological studies.

Hybridization based techniques for detection of *Mycoplasmas* include the use of DNA probes. By way of example, radiolabelled oligonucleotide probes have been described that target the *M. genitalium* 16S rRNA.

Because of the problems encountered using traditional procedures for the diagnosis of *M. genitalium* infection, such as culture and serology, knowledge about its pathogenicity has been slow to accrue. Extensive clinical studies have only become possible with the advent of PCR-based assays.

PCR has proven a particularly useful tool for detecting fastidious *Mycoplasmas* due to its very high sensitivity (Jensen, Dan. Med. Bull. 2006; 53: 1-27). However, until recently, a major drawback of this technique has been the lack of commercial kits.

PCR assays have recently been developed that amplify and detect the *M. genitalium* MgPa gene (eg. Jensen *et al.*, J. Clinical Microbiology, Feb. 2004, pages 683-692). The MgPa gene encodes the major *M. genitalium* adhesion protein responsible for attachment to host cells, in particular for attachment to the epithelial cells, such as those of the human fallopian tube.

Evasion of the host immune system by antigenic variation of surface components such as adhesins enables *Mycoplasma* such as *M. genitalium* to adapt to changing environments and selection pressures. This antigenic variation leads to considerable sequence variation in genes that encode

surface proteins such as the MgPa gene. Thus, diagnostic assays that rely on detection of surface proteins and corresponding nucleic acids fail to detect some variants, resulting in poor sensitivity of the assay.

Other published PCR-based assays for detecting *M. genitalium* are directed towards the 16S rRNA gene (Yoshida *et al.*, J. Clin. Microbiol. 2002: 40:1451-1455). However, the specificity and sensitivity of this assay has been put into question due to the high level of homology between the M. *pneumoniae* and *M. genitalium* 16S rRNA genes, and the preponderance of secondary structures in the 16S rRNA.

A LightCycler assay has also been described, directed towards detection of the P115 (MG299) gene with locked nucleic acid probes (Dupin *et al.*, Clin. Infect. Dis. 2003; 37:602-605).

There is, therefore, a need to provide an alternative and/ or improved system for detecting *M. genitalium*.

As used herein, except where the context requires otherwise, the term "comprise" and variations of the term, such as "comprising", "comprises" and "comprised", are not intended to exclude other additives, components, integers or steps.

Reference to any prior art in the specification is not, and should not be taken as, an acknowledgment, or any form of suggestion, that this prior art forms part of the common general knowledge in Australia or any other jurisdiction or that this prior art could reasonably be expected to be ascertained, understood and regarded as relevant by a person skilled in the art.

According to a first aspect, the present invention provides a method for detecting *M. genitalium* nucleic acid in a sample, comprising: (i) amplifying a nucleic acid sequence comprising SEQ ID NO: 1 or a fragment thereof; and (ii) detecting said amplified nucleic acid sequence.

According to a second aspect, the present invention provides a method for detecting *M. genitalium* nucleic acid in a sample, comprising: (i) contacting said sample with a

probe, wherein the probe binds to a target sequence within SEQ ID NO: 1, or the complement thereof; and (ii) detecting binding of said probe to said target site.

The present invention provides improved specificity and sensitivity compared to existing PCR-based tests for *M. genitalium*.

The *M. genitalium* Mg219 gene (SEQ ID NO: 1) is 447 nucleotides long and encodes a 148 amino acid polypeptide of unknown function (SEQ ID NO: 2).

The Mg219 gene (SEQ ID NO: 1) is located from nucleotide residue 265596 to nucleotide residue 266042 of the complete genome of *M. genitalium* (as detailed in Accession No. NC_000908, *M. genitalium* G-37).

Mg219 is found within an operon consisting of the Mg217, Mg218 and Mg219 genes, and transcription of Mg219 is continuous from Mg217 and Mg218 (Musatovova O. *et al.*, FEMS Microbiol. Lett., 2003, 5, 229(1) pages 73-81). Mg218 encodes a high molecular mass cytadherence-related protein required for tip-mediated adherence. The function of Mg219 remains unknown, but due to the proximity to Mg218 in the *M. genitalium* genome, it is possible that Mg219 may be involved in adherence.

A detailed review by the present Applicant of all published *Mycoplasma* whole genome sequences revealed that the Mg219 gene has no homologues in *Mycoplasma* species other than *M. genitalium*.

Furthermore, by searching published genetic and protein databases, the present Applicant was unable to identify any gene having significant homology to the Mg219 gene, or any protein having significant homology to the encoded Mg219 polypeptide.

These findings were confirmed using the ORFANGE web-based program, which identifies genes that are only found once in a bacterial genome and not in other genomes.

Thus, the present Applicant has unexpectedly identified that the Mg219 gene is specific to *M. genitalium*, and is thus useful for identification of *M. genitalium* in a sample.

It is particularly important to keep the incidence of "false positive" results as low as possible in the field of detecting and diagnosing sexually transmitted diseases, due to the sensitive nature of the test results for the patient, cost

and implications of inappropriate treatment, and the legal implications of providing an incorrect result.

Advantageously, because the Mg219 gene has no sequence homologues in other *Mycoplasma* species, or in other published sequence databases, the detection assay of the present invention, which is based on detection of Mg219 nucleic acid sequences, is very specific, with a very low incidence of "false positive" results.

As illustrated in the Examples (below), when a detection assay according to one embodiment of the present invention was tested by the present Applicant against DNA from all known human Mollicutes, some other bacterial species and pathogenic micro-organisms, the assay only detected *M. genitalium* – thus illustrating the high degree of specificity of the assay.

The target Mg219 gene detected in the present invention is also an advantageous target gene because it is conserved between different strains of *M. genitalium*. Thus, the assay of the present invention advantageously detects all known strains of *M. genitalium* that have been tested in the assay to date.

A sample may be for instance, a food, sewerage, environmental, veterinary or clinical sample. In one aspect, the method may be used for detection of *M. genitalium* in a clinical sample.

Clinical samples may include urethral swabs, vaginal swabs, cervical swabs, rectal swabs, penile swabs, throat/oral swabs, urine, blood, respiratory tract samples, synovial fluids, cerebro-spinal fluid, liquid based cytology samples, tissue biopsies, ulcer samples, conjunctivitis samples and any other samples from animals, particularly from humans. In one aspect, the sample may comprise semen or eggs.

Thus, in one aspect, the method of the present invention comprises the step of amplifying *M. genitalium* nucleic acid.

In this application, the expressions "amplified nucleic acid sequence" and "amplicon" are used interchangeably and have the same meaning.

Specifically, in one aspect, the method of the present invention comprises amplifying a nucleic acid sequence comprising SEQ ID NO: 1. The amplified nucleic acid sequence (amplicon) may consist of SEQ ID NO: 1.

In one aspect, the method of the present invention comprises amplifying a nucleic acid sequence comprising a fragment of SEQ ID NO: 1. The amplified nucleic acid sequence (amplicon) may consist of a fragment of SEQ ID NO: 1.

Alternatively, the amplicon may comprise a fragment of SEQ ID NO: 1 and also 1 or more additional nucleotides that are located upstream or downstream of SEQ ID NO: 1 in the *M. genitalium* genome.

A fragment of SEQ ID NO: 1 is preferably at least 10 consecutive nucleotides of SEQ ID NO: 1, and is more preferably at least 25 nucleotides, more preferably at least 50 nucleotides, and may be at least 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375 or 400 nucleotides of SEQ ID NO: 1. A fragment of SEQ ID NO: 1 is preferably up to 440 consecutive nucleotides of SEQ ID NO: 1, more preferably up to 425 nucleotides, more preferably up to 400 nucleotides, and may be up to 375, 350, 325, 300, 275, 250, 225, 200, 175, 150, 125, 100, 75, 50, or 25 nucleotides of SEQ ID NO: 1.

In a specific example, the amplicon may comprise a stretch of consecutive nucleotides spanning the nucleotide located 11 positions upstream of SEQ ID NO: 1 in the *M. genitalium* genome (ie. nucleotide position -11) to the nucleotide at position 240 of SEQ ID NO: 1. Thus, in this specific example, the amplicon comprises a 240 nucleotide fragment of SEQ ID NO: 1, from nucleotide 1 of SEQ ID NO: 1 to nucleotide 240 of SEQ ID NO: 1 (plus the 11 nucleotides upstream of SEQ ID NO: 1 in the *M. genitalium* genome).

Amplification may be carried out by methods known in the art, preferably by PCR. Examples of PCR platforms suitable for conducting the amplification step of the present invention include real-time platforms such as Rotor-gene, LightCycler and Taqman platforms.

In one aspect, amplification of *M. genitalium* nucleic acid is carried out using a pair of sequence specific oligonucleotide primers, wherein said primers bind to target sites in the *M. genitalium* nucleic acid. Under suitable conditions, the primers are extended, resulting in nucleic acid synthesis. A skilled person would be able to determine suitable conditions for promoting amplification of a nucleic acid sequence comprising a fragment of SEQ ID NO: 1.

For the avoidance of doubt, in the context of the present invention, the definition of an oligonucleotide primer does not include the full length Mg219 gene (or complement thereof).

Primers of the present invention are designed to bind to the target gene sequence based on the selection of desired parameters, using conventional software, such as Primer Express (Applied Biosystems). In this regard, it is preferred that the binding conditions are such that a high level of specificity is provided. The melting temperature (Tm) of the primers is preferably in excess of 50°C and is most preferably about 60°C. A primer of the present invention preferably binds to target *M. genitalium* nucleic acid but is preferably screened to minimise self-complementarity and dimer formation (primer-to-primer binding).

The primer pair comprises forward and reverse oligonucleotide primers.

A forward primer binds to the complementary (ie. anti-sense) strand of the target *M. genitalium* nucleic acid and a reverse primer binds to the coding (sense) strand of the target *M. genitalium* nucleic acid.

The forward and reverse oligonucleotide primers are typically at least 5 nucleotides long, preferably at least 10 nucleotides long, more preferably at

least 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides long. Preferably, the primers are up to 60 nucleotides long, preferably up to 55 nucleotides long, more preferably up to 50, 45, 40, 35, 30, 25, 20, 15 or 10 nucleotides long.

In one aspect, the forward primer and/ or the reverse primer is about 50-55 nucleotides long. In one aspect, the forward primer and/ or the reverse primer is about 20-25 nucleotides long. In one aspect, the forward primer and/ or the reverse primer is about 25-30 nucleotides long. In one aspect, the forward primer is about 29 nucleotides long and the reverse primer is about 27 nucleotides long. It is an advantage to use shorter primers, as this enables faster annealing to target nucleic acid.

The forward primer binds to a target site within the *M. genitalium* nucleic acid. This target site may be located within the complement of SEQ ID NO: 1 – ie. the target site may be located between nucleotide 1 and nucleotide 447 of the complement of SEQ ID NO: 1.

Reference to "a target site between (for example) nucleotide residues 1 and 447 of the complement of SEQ ID NO: 1" does not imply that the target site extends between or is defined by the recited residues. Thus, the target site for the forward primer may be 447 nucleotides long, or may be shorter, even considerably shorter than 447 nucleotides in length (eg. up to about 200, 100, 75, 60, 50, 40, 30, 25, 20, 15, 10 nucleotides long), so long as it is located somewhere within the recited range of nucleotide residues.

In one aspect, the target site of the forward primer is about 50-55 nucleotides long. In one aspect, the target site of the forward primer is about 25-30 nucleotides long, preferably about 29 nucleotides long.

Alternatively, the target site for the forward primer within the *M. genitalium* nucleic acid may be located within upstream nucleic acid sequences that flank the complement of SEQ ID NO: 1 in the *M. genitalium* genome.

SEQ ID NO: 1 is located from nucleotide residue 265596 to nucleotide residue 266042 of the complete genome of *M. genitalium* (as detailed in Accession No. NC_000908, *M. genitalium* G-37). Thus, the nucleic acid sequences that are located upstream of the complement of SEQ ID NO: 1 in the *M. genitalium* genome have nucleotide numbers 1-265595.

Alternatively, the forward primer may bind to a target site that overlaps the complement of SEQ ID NO: 1 and upstream nucleic acid sequences that flank the complement of SEQ ID NO: 1 in the *M. genitalium* genome.

Upstream nucleotides are typically designated using negative numbers. By way of example, the nucleotide residue 11 residues upstream from the start of the complement of SEQ ID NO: 1 in the *M. genitalium* genome (ie. nucleotide residue 265585 of the complete *M. genitalium* genome) is designated "nucleotide -11".

The nucleic acid sequence of the 60 nucleotide residues located upstream from SEQ ID NO: 1 in the *M. genitalium* genome (ie. nucleotides –60 to –1 of SEQ ID NO: 1) are represented by SEQ ID NO: 6. Thus, nucleotide residue number 1 of SEQ ID NO: 6 is nucleotide residue number –60 of SEQ ID NO: 1, and nucleotide residue number 60 of SEQ ID NO: 6 is nucleotide residue number –1 of SEQ ID NO: 1.

In one aspect, the forward primer binds to a target site located between nucleotides –80 to +450 of a nucleic acid strand complementary to SEQ ID NO: 1. Within this range of nucleotide residues, the target site for the forward primer is preferably located from nucleotide residue –76 of a nucleic acid strand complementary to SEQ ID NO: 1, and may be located from nucleotide residue –60, –55, –53, –50, –25, –10, 1, 25, 51, 101, 151, 201, 251, 301, 351 or 401 of a nucleic acid strand complementary to SEQ ID NO: 1. Preferably, the target site for the forward primer is located up to nucleotide residue 447 of a nucleic acid strand complementary to SEQ ID NO: 1, and may be located up to nucleotide residue 425, 400, 350, 300, 250, 200, 150, 100, 50, 25, 1, –5, –10, –25, –50 or –55 of a nucleic acid strand complementary to SEQ ID NO: 1.

In one aspect, the forward primer binds to a target site located between nucleotide residues –25 to 25 of a nucleic acid strand complementary to SEQ ID NO: 1. Preferably, the target site for the forward primer is located from nucleotide residue –20, –15, –14, –13 or –12 and up to nucleotide residue 22, 20 or 19 of a nucleic acid strand complementary to SEQ ID NO: 1.

In one aspect, the target site for the forward primer is defined by nucleotide residues –11 to 18 of a nucleic acid strand complementary to SEQ ID NO: 1. This target region consists of nucleotide residues 50-60 of a nucleic acid strand complementary to SEQ ID NO: 6, followed by nucleotide residues 1-18 of a nucleic acid strand complementary to SEQ ID NO: 1.

For the avoidance of any doubt, the above numbering system applied to the nucleic acid residues of the complementary strand of SEQ ID NO: 1 (and upstream sequences, SEQ ID NO: 6) is based on the numbering of the nucleic acids of SEQ ID NO: 1 (and upstream sequences, SEQ ID NO: 6) to which they are complementary.

The reverse primer binds to a target site within the *M. genitalium* nucleic acid. This target site may be located within SEQ ID NO: 1 – ie. the target site may be located between nucleotide 1 and nucleotide 447 of SEQ ID NO: 1.

Reference to "a target site between (for example) nucleotide residues 1 and 447 of SEQ ID NO: 1" does not imply that the target site extends between or is defined by the recited residues. Thus, the target site for the reverse primer may be 447 nucleotides long, or may be shorter, even considerably shorter than 447 nucleotides in length (eg. up to 200, 100, 75, 60, 50, 40, 30, 25, 20, 15 nucleotides long), so long as it is located somewhere within the recited range of nucleotide residues.

In one aspect, the target site for the reverse primer is about 50-55 nucleotides long. In one aspect, the target site for the reverse primer is about 25-30 nucleotides long, preferably about 27 nucleotides long.

Alternatively, the target site for the reverse primer within the *M. genitalium* nucleic acid may be located within downstream nucleic acid sequences that flank SEQ ID NO: 1 in the *M. genitalium* genome.

SEQ ID NO: 1 is located from nucleotide residue 265596 to nucleotide residue 266042 of the complete genome of *M. genitalium* (as detailed in Accession No. NC_000908, *M. genitalium* G-37). Thus, the nucleic acid sequences that are located downstream of SEQ ID NO: 1 in the *M. genitalium* genome have nucleotide numbers 266043 to 580074.

Alternatively, the reverse primer may bind to a target site that overlaps SEQ ID NO: 1 and downstream nucleic acid sequences that flank SEQ ID NO: 1 in the *M. genitalium* genome.

Downstream nucleotides are designated with positive numbers, continuing from the number of the last nucleotide of SEQ ID NO: 1 (residue 447). By way of example, the nucleotide residue located 20 residues downstream from the end of SEQ ID NO: 1 in the *M. genitalium* genome (ie. nucleotide residue 266063 of the complete *M. genitalium* genome) is designated "nucleotide +467" (ie. 447 +20 = 467).

The nucleic acid sequence of the 53 nucleotide residues located downstream from SEQ ID NO: 1 in the *M. genitalium* genome (ie. nucleotides +448 to +500 of SEQ ID NO: 1) are represented by SEQ ID NO: 7. Thus, nucleotide residue number 1 of SEQ ID NO: 7 is nucleotide residue number +448 of SEQ ID NO: 1, and nucleotide residue number 53 of SEQ ID NO: 7 is nucleotide residue number +500 of SEQ ID NO: 1.

In one aspect of the invention, the reverse primer binds to a target site between nucleotide residues –5 to +525 of SEQ ID NO: 1. Within this range of nucleotide residues, the target site for the reverse primer is preferably located from nucleotide residue –3 of SEQ ID NO: 1, and may be located from nucleotide residue 1, 10, 25, 48, 98, 148, 198, 248, 298, 348, 398, +448 or

+503 of SEQ ID NO: 1. Preferably, within this range of nucleotide residues, the target site for the reverse primer is located up to nucleotide residue +502 of SEQ ID NO: 1, and may be located up to nucleotide residue 447, 397, 347, 297, 247, 197, 147, 97, 47 or 25 of SEQ ID NO: 1.

In one aspect, the reverse primer binds a target site between nucleotide residues 200 to 250 of SEQ ID NO: 1. Within the range of nucleotide residues, the target site for the reverse primer is preferably located from nucleotide residue 205, 210, 211, 212 or 213 of SEQ ID NO: 1 and up to nucleotide residue 245, 244, 243, 242 or 241 of SEQ ID NO: 1. In one aspect, the reverse primer binds to a target region defined by residues 214 to 240 of SEQ ID NO: 1.

The amplified nucleic acid sequence (amplicon) is preferably at least 10 nucleotides long, more preferably at least 20 nucleotides long, more preferably at least 40 nucleotides long, and may be at least 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375 or 400 nucleotides long. The amplicon is preferably up to 440 nucleotides long, preferably up to 425 nucleotides long, more preferably up to 400 nucleotides long, and may be up to 375, 350, 325, 300, 275, 250, 225, 200, 175, 150, 125, 100, 75, 50, 40 or 20 nucleotides long. In one aspect, the amplicon is about 225-275 nucleotides long, preferably about 251 nucleotides long.

Particularly good results have been obtained using a forward primer selected from SEQ ID NOs: 4, 8 or 31, as shown in the table below, which bind to a target site defined by nucleotide residues –11 to 18 of a nucleic acid strand complementary to SEQ ID NO: 1.

SEQ ID NO:	SEQUENCE
4	5' GAC AGT TCA TTA TGC GCA CCA GTT ACT TG 3'
8	5' CAT AGT TCA TTA TGC GCA CCA GTT ACT TG 3'
31	5' CAT AGT TCA TTA TGC ACA CCA GTT ACT TG 3'

Particularly good results have been obtained using a forward primer selected from SEQ ID NOs: 9-19, as shown in the table below (see also Example 4).

9	5' CCA CTT AAC TTT ATT ACC CGT CC 3'
10	5' TGT TTT CAA AAG TAA TTT GCC ACC GAA ACT AAG TAA GGA TGA CAT AGT TCA TT 3'
11	5' ATG CGC ACC AGT TAC TTG AAA AAA ATA CCC ATA ATG AAT AGT GAT AGT GA 3'
12	5' TCT AAA ACT CCA AAA GGT GTG GAT CGA GCG GCA TGT TGA TCA AGA TGA AC 3'
13	5' TTA GTT TAA CAA CTA CTG CAG TTG AAC TTA AAA AGA GTG ATG AAC AAA AA 3'
14	5' CCT GTT GCC ATT AAA AGT AGT GAC TTT ATT GGT CAT GAA GAG TTA ATC TC 3'
15	5' TGT TCC AGT TTT ACT AAT CCC AAC CCC TGT TGT TAA AGA GAT TGA TCA AC 3'
16	5' CAG CAG TTA TTC CTC CAG TTA AAG CAA AAC CAA AAG CAA CTA AAA AGA AA 3'
17	5' ACT CCT GTT AAA TCA AAA CCA ACT AGT AAA TCA ACT AAA CAA ACA AAA CC 3'
18	5' TAA ACA ATC CAA GCC CAA ATC AAA ACA AGT TCA ACA AAC CAA AGC TAA AC 3'
19	5' CAA CCC AAA TTC AAA CAA AAA AAA GCA ATA AAA AAA CCA GAT CTT AAT CT 3'

Particularly good results have been obtained using a reverse primer of SEQ ID NO: 5, as shown in the table below, which binds to a target site defined by nucleotide residues 214 to 240 of SEQ ID NO: 1.

SEQ ID NO:	SEQUENCE
5	5' CTC TTT AAC AAC AGG GGT TGG GAT TAG 3'

Particularly good results have been obtained using a reverse primer selected from SEQ ID NOs: 20-30, as shown in the table below (see also Example 4).

20	5' GAT TAA CCC CAG GTA GTT CTT CC 3'
21	5' TAT TCC TTT CCA GTT TTT AGT TAA AAC TAC TGT TGT TAA CAC TAA AAA AAC CAG A 3'
22	5' TTA AGA TCT GGT TTT TTT ATT GCT TTT TTT TGT TTG AAT TTG GGT TGG TT 3'
23	5' TAG CTT TGG TTT GTT GAA CTT GTT TTG ATT TGG GCT TGG ATT GTT TAG GT 3'
24	5' TTT GTT TGT TTA GTT GAT TTA CTA GTT GGT TTT GAT TTA ACA GGA GTT TT 3'
25	5' CTT TTT AGT TGC TTT TGG TTT TGC TTT AAC TGG AGG AAT AAC TGC TGG TT 3'
26	5' GAT CAA TCT CTT TAA CAA CAG GGG TTG GGA TTA GTA AAA CTG GAA CAG AG 3'
27	5' ATT AAC TCT TCA TGA CCA ATA AAG TCA CTA CTT TTA ATG GCA ACA GGT TT 3'
28	5' TTG TTC ATC ACT CTT TTT AAG TTC AAC TGC AGT AGT TGT TAA ACT AAG TT 3'
29	5' CAT CTT GAT CAA CAT GCC GCT CGA TCC ACA CCT TTT GGA GTT TTA GAT CA 3'
30	5' CTA TCA CTA TTC ATT ATG GGT ATT TTT TTC AAG TAA CTG GTG CGC ATA AT 3'

It will, however, be appreciated that variants may be employed, which differ from the above-mentioned primer sequences by one or more nucleotides. In this regard, conservative substitutions are preferred.

Thus, in one aspect, the forward primer comprises a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, more preferably at least 97%, most preferably 100% sequence identity to a sequence selected from SEQ ID NOs: 4, 8-19 or 31.

Thus, in one aspect, the reverse primer comprises a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, more preferably at least 97%, most preferably 100% sequence identity to a sequence selected from SEQ ID NO: 5 or SEQ ID NOs: 20-30.

Preferably, the forward primer consists of a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, more preferably at least 97%, most preferably 100% sequence identity to a sequence selected from SEQ ID NO: 4, 8-19 or 31.

Preferably, the reverse primer consists of a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, more preferably at least 97%, most preferably 100% sequence identity to a sequence selected from SEQ. ID NOs: 5 or 20-30.

Fragments of the above-mentioned primer sequences may also be employed.

Preferably, a fragment of forward primers SEQ ID NOs: 4, 8 and 31 comprises at least 15 consecutive nucleotides of said primer sequences, more preferably at least 20, 21, 22, 23, 24, 25, 26 or 27 consecutive nucleotides of said primer sequences, and may comprise up to 28 consecutive nucleotides of said primer sequences.

Preferably, a fragment of reverse primer SEQ ID NO: 5 comprises at least 15 consecutive nucleotides of said primer sequence, more preferably at least 20, 21, 22, 23, 24 or 25 consecutive nucleotides of said primer sequence, and may comprise up to 26 consecutive nucleotides of said primer sequence.

Preferably, a fragment of primer SEQ ID NOs: 9 and 20 comprises at least 10 consecutive nucleotides of said primer sequences, more preferably at least 15, 16, 17, 18, 19, 20 or 21 consecutive nucleotides of said primer sequences, and may comprise up to 22 consecutive nucleotides of said primer sequences.

Preferably, a fragment of primer SEQ ID NOs: 10-30 comprises at least 30 consecutive nucleotides of said primer sequences, more preferably at least 35, 40, 41, 42, 43, 44, 45, 46, 47 or 48 consecutive nucleotides of said primer sequences, and may comprise up to 49 consecutive nucleotides of said primer sequences.

In the method of the present invention, any forward primer selected from SEQ ID NOs: 4, 8-19 or 31 (or fragments thereof, or variants thereof having a % identity thereto as discussed above) may be used in combination with any reverse primer selected from SEQ ID NOs: 5 or 20-30 (or fragments thereof, or variants thereof having a % identity thereto as discussed above).

In one aspect, the method may employ forward primer SEQ ID NO: 4, 8 or 31 and reverse primer SEQ ID NO: 5 (or fragments thereof, or variants thereof having % identity thereto as discussed above).

In one aspect, the method may employ forward primer SEQ ID NO: 9 and reverse primer SEQ ID NO: 20 (or fragments thereof, or variants thereof having % identity thereto as discussed above).

It is an option for at least one of the primers to comprise a minor groove binder (MGB) component.

The detection step may be carried out by any known means. In one aspect, the amplified nucleic acid sequence is detected by a method comprising gel electrophoresis.

Alternatively, or in addition, the step of detecting the amplified nucleic acid sequence may comprise contacting said amplified nucleic acid sequence with a probe, wherein the probe binds to a target site within said amplified nucleic acid sequence, or the complement thereof, and detecting binding of said probe to said amplified nucleic acid sequence.

Suitable probes for use in the methods of the present invention are ligands that bind specifically to *M. genitalium* nucleic acid. Such ligands may be oligonucleotide ligands or protein ligands, for example, antibodies. However, it is preferred that the probes are oligonucleotide probes.

For the avoidance of doubt, in the context of the present invention, the definition of an oligonucleotide probe does not include the full length Mg219 gene (or complement thereof).

Probes are designed to bind to the target gene sequence (ie. within the amplicon, or within SEQ ID NO: 1) based on a selection of desired parameters, using conventional software. It is preferred that the binding conditions are such that a high level of specificity is provided – ie. binding occurs under "stringent conditions". In general, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target

sequence binds to a perfectly matched probe. In this regard, the T_m of probes of the present invention, at a salt concentration of about 0.02M or less at pH 7, is preferably above 60°C, more preferably about 70°C. Premixed binding solutions are available (eg. EXPRESSHYB Hybridisation Solution from CLONTECH Laboratories, Inc.), and binding can be performed according to the manufacturer's instructions. Alternatively, a person skilled in the art can devise variations of these binding conditions.

Following binding, washing under stringent (preferably highly stringent) conditions removes unbound nucleic acid molecules. Typical stringent washing conditions include washing in a solution of 0.5-2x SSC with 0.1% SDS at 55-65°C. Typical highly stringent washing conditions include washing in a solution of 0.1-0.2x SSC with 0.1%SDS at 55-65°C. A skilled person can readily devise equivalent conditions for example, by substituting SSPE for the SSC in the wash solution.

It is preferable to screen the probes to minimise self-complementarity and dimer formation (probe-probe binding). Preferred probes of the present invention are selected so as to have minimal homology with human DNA. The selection process may involve comparing a candidate probe sequence with human DNA and rejecting the probe if the homology is greater than 50%. The aim of this selection process is to reduce annealing of probe to contaminating human DNA sequences and hence allow improved specificity of the assay.

In one embodiment, conjugation of a minor groove binder (MGB) to the probe stabilises nucleic acid duplexes, causing a desirable increase in oligonucleotide Tm (the temperature at which the duplex separates). This increase in Tm enables considerably shorter probes to be used than would otherwise be possible.

In one aspect, the oligonucleotide probe is at least 5 nucleotides long, more preferably at least 10 nucleotides long, more preferably at least 12, 13, 14, 15

or 16 nucleotides long. Preferably, the probe is up to 50 nucleotides long, more preferably up to 30 nucleotides long, and more preferably up to 20, 19 or 18 nucleotides long. It is an advantage to use shorter probes, as this enables faster annealing to target *M. genitalium* nucleic acid.

The complement of a nucleic acid sequence binds via complementary base-pairing to the nucleic acid sequence. In the present invention, a "complementary strand" means the anti-sense nucleic acid strand, which binds via complementary base-pairing to a sense strand.

Thus, in one aspect, the probe binds to a target sequence within the coding (sense) strand of the target *M. genitalium* nucleic acid (ie. within SEQ ID NO: 1). In an alternative aspect, the probe binds to a target sequence within the complementary, non-coding (anti-sense) strand of the target *M. genitalium* nucleic acid (ie. within the complement of SEQ ID NO: 1).

The target site to which the probe binds may be located anywhere within SEQ ID NO: 1, or within the complement of SEQ ID NO: 1. If the nucleic acid in the sample has previously been amplified, the probe binds to a target site within said amplified nucleic acid sequence, or the complement thereof.

Thus, in one aspect, the probe binds to a target site located anywhere between residues 1 and 447 of SEQ ID NO: 1, or the complement thereof. In this regard, reference to "a target site between (for example) residues 1 and 447 of SEQ ID NO: 1 or the complement thereof" does not imply that the target site extends between or is defined by the recited residues. Thus, the target site may be 447 nucleotides long, or may be shorter, even considerably shorter than 447 nucleotides in length (eg. up to 200, 100, 75, 60, 50, 40, 30, 25, 20, 15 or 10 nucleotides long).

Thus, in one aspect, the probe binds to a target site between nucleic acid residues 1-447 of SEQ ID NO: 1 or the complement thereof. Within this range of nucleotide residues, the target site for the probe is preferably located from nucleotide residue 10 of SEQ ID NO: 1 or the complement thereof, more

preferably from nucleotide residue 15 of SEQ ID NO: 1 or the complement thereof, and may be located from nucleotide residue 25, 50, 75, 100, 150, 200, 250, 300, 350 or 400 of SEQ ID NO: 1 or the complement thereof. Within this range of nucleotide residues, the target site for the probe is preferably located up to nucleotide residue 440, more preferably up to nucleotide residue 420 of SEQ ID NO: 1 or the complement thereof, and may be located up to residue 400, 350, 300, 250, 200, 150, 100 or 50 of SEQ ID NO: 1 or the complement thereof.

In one aspect, the probe binds to a target region located between nucleotide residues 50-100 of the nucleic acid strand complementary to SEQ ID NO: 1. Preferably, the target site for the probe is located from nucleotide residue 55, 60, 63, 64 or 65 and up to nucleotide residue 90, 85, 84 or 83 of the nucleic acid strand complementary to SEQ ID NO: 1. Most preferably, the target site for the probe is defined by residues 66-82 of the nucleic acid strand complementary to SEQ ID NO: 1.

For the avoidance of any doubt, the above numbering system applied to the nucleic acid residues of the complementary strand of SEQ ID NO: 1 is based on the numbering of the nucleic acids of SEQ ID NO: 1 to which they are complementary.

Good results have been obtained using a probe consisting of the nucleic acid sequence SEQ ID NO: 3, as shown in the table below.

SEQ ID NO:	SEQUENCE
3	5'-GGT GTG GAT CGA GCG GC -3'

It will, however, be appreciated that variants may be employed, which differ from the above-mentioned probe sequence by one or more nucleotides. In this regard, conservative substitutions are preferred.

Thus, in one aspect, the probe comprises a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, most preferably 100% sequence identity to SEQ ID NO: 3.

In one aspect, the probe consists of a nucleic acid sequence having at least 60%, preferably at least 65%, preferably at least 70%, preferably at least 75%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, most preferably 100% sequence identity to SEQ ID NO: 3.

An alternative means for defining variant probe sequences is by defining the number of nucleotides that differ between the variant sequence and the probe sequence of the present invention. In this regard, the present invention embraces probe sequences that differ from SEQ ID NO: 3 by no more than 6 nucleotides, preferably by no more than 5 nucleotides, preferably by no more than 4 nucleotides, more preferably by no more than 3 nucleotides, yet more preferably by no more than 2 nucleotides, and most preferably by no more than 1 nucleotide.

A fragment of the above-mentioned probe sequence may also be employed, wherein the fragment comprises at least 10, preferably at least 11, 12, 13, 14, 15 or 16 consecutive nucleotides of SEQ ID NO: 3.

In one aspect, the probe may be immobilised onto a solid support or platform. The support may be a rigid solid support made from, for example, glass or plastic, or else the support may be a nylon or nitrocellulose membrane, or other membrane. 3D matrices are suitable supports for use with the present invention - eg. polyacrylamide or PEG gels. In one embodiment, the solid support may be in the form of beads, which may be sorted by size or fluorophores.

The probes may be immobilised to the solid support by a variety of means. By way of example, probes may be immobilised onto a nylon membrane by UV cross-linking. Biotin-labelled probes may be bound to streptavidin-coated substrates, and probes prepared with amino linkers may be immobilised onto silanised surfaces.

Another means of immobilising probe is via a poly-T tail, preferably at the 3' end. The poly-T tail consists of a run of from 1 to 100 thymine residues added to the probe at the 3' end with a terminal transferase. Preferably, from 1 to 20 thymine residues are added. The poly-T tail is then baked or UV cross-linked onto the solid substrate. Addition of a poly-T tail appears to have two functions. First, the poly-T tail increases the amount of probe that is immobilised onto the solid support. Second, the poly-T tail conforms the probe in such a way as to improve the efficiency of hybridisation.

It is an option for the probe to comprise a minor groove binder (MGB) component.

In one aspect, binding of probe to *M. genitalium* nucleic acid provides a detectable signal, which may be detected by known means. A detectable signal may be, for example, a radioactive signal or a fluorescent signal, such as a change in fluorescence.

In one embodiment, the probe is labelled and the assay comprises detecting the label and correlating presence of label with presence of *M. genitalium* nucleic acid. The label may be a radiolabel but is preferably non-radioactive, such as a fluorescent label. By way of example, the label may be digoxygenin or fluorescein-isothiocyanate (FITC).

The label may be detected directly, such as by exposure to photographic or X-ray film, or indirectly, for example, in a two-phase system. An example of indirect label detection is binding of an antibody to the label. In another aspect, the probe is labelled with biotin and is detected using streptavidin

bound to a detectable molecule or to an enzyme, which generates a detectable signal.

In one aspect, prior to detecting binding of the probe to its target sequence, the method comprises amplifying *M. genitalium* nucleic acid. It may be desirable to amplify the target human papillomavirus nucleic acid if the sample is small and/ or comprises a heterogeneous collection of DNA sequences.

Most preferably, the amplification step is carried out prior to contacting the nucleic acid with the probe.

Amplification of *M. genitalium* nucleic acid is preferably carried out as described above. In this regard, the amplification is preferably carried out by contacting the sample with forward and reverse oligonucleotide primers as described above. Said primers bind to target sites in the *M. genitalium* nucleic acid, under conditions suitable to promote amplification of a nucleic acid sequence comprising at least the target sequence to which the probe binds within SEQ ID NO: 1.

The present invention enables quantitative estimates of pathogen load to be determined. Determining pathogen load has many useful applications, such as for clinical guidance and for determining therapy.

Thus, in one aspect, the present invention provides a method of quantitating *M. genitalium* pathogen load in a sample of interest, comprising: (a) carrying out a detection method according to the present invention on said sample of interest; and (b) carrying out said method on a test sample of predetermined known *M. genitalium* pathogen load; and (c) comparing the signal detected from the sample of interest with the signal detected from the test sample; and thereby quantitating *M. genitalium* pathogen load in the sample of interest.

In another aspect, the method of the present invention is useful for determining efficacy of a course of treatment over a period of time, for example a course of drug therapy.

Thus, in one aspect, the present invention provides an *in vitro* method of determining drug efficacy over the course of a period of drug therapy, comprising: (a) carrying out the present method on a first sample obtained at a first time point within or prior to the period of drug therapy; (b) carrying out the present method on a sample at one or more later time points within the period of drug therapy; and (c) comparing the signal detected from the first sample with the signal detected from the one or more later samples; and thereby determining drug efficacy over the course of the period of drug therapy.

In one aspect, the method of the present invention is useful for detecting and/ or monitoring the development of resistance to a drug (eg. an anti-microbial such as an antibiotic for *M. genitalium*) over a period of time.

Thus, in one aspect, the present invention provides an *in vitro* method of detecting and/ or monitoring the development of resistance of *M. genitalium* to a drug, comprising (a) carrying out the present method on a first sample obtained at a first time point within or prior to exposure of the *M. genitalium* to the drug; (b) carrying out the present method on a sample at one or more later time points following exposure of the *M. genitalium* to the drug; and (c) comparing the signal detected from the first sample with the signal detected from the one or more later samples; wherein the absence of a reduction in the signal, or an increase in the signal, indicates that the *M. genitalium* have developed resistance to the drug.

In one aspect, the method of the present invention is useful for screening human semen and/ or human eggs prior to artificial insemination, to confirm the presence or absence of *M. genitalium*. In this regard, confirmation of the absence of *M. genitalium* infection may be required prior to IVF treatment. Hence, the method of the present invention advantageously reduces (and preferably substantially eliminates) the risk of transmitting *M. genitalium* to a patient undergoing *in vitro* fertilisation (IVF) treatment.

Thus, in one aspect, the present invention provides an *in vitro* method of screening a sample comprising human semen and/ or eggs for *M. genitalium*, comprising carrying out the present detection method; wherein the absence of a signal indicates that the sample is free of *M. genitalium*.

The detection method of the present invention may also be useful for screening people undergoing IVF treatment, and surrogate mothers, for *M. genitalium* infection.

The detection method of the present invention may also be useful for occupational screening for *M. genitalium* infection. The detection method of the present invention may also be useful for forensic testing – eg. in rap, sexual assault or child abuse cases.

The invention also provides reagents for use in the above-described methods of the present invention.

Hence, the present invention provides a forward primer as described above for use in accordance with the invention. The present invention also provides a reverse primer as described above for use in accordance with the invention. The present invention also provides a pair of forward and reverse oligonucleotide primers selected from the primers described above in accordance with the invention.

The present invention provides a probe as described above for use in accordance with the present invention.

Also provided by the present invention is a kit for detecting *M. genitalium* nucleic acid, comprising a forward primer as described above. Optionally, the kit includes a reverse primer as described above. Optionally, the kit includes a probe as described above.

The present invention is discussed in more detail by means of the Examples described below, and by the Figures.

Figure 1 illustrates the relationship between the Mg219 gene (SEQ ID NO: 1), probe SEQ ID NO: 3, forward primer SEQ ID NO: 4, 8 or 31, and reverse primer SEQ ID NO: 5.

In more detail, Figure 1 illustrates the 447 nucleotide sequence SEQ ID NO: 1 with 3 superimposed shaded regions representing the target/ binding sites for the probe SEQ ID NO: 3, the forward primer SEQ ID NO: 4, 8 or 31, and the reverse primer SEQ ID NO: 5. In this regard, although the probe SEQ ID NO: 3 and the forward primer SEQ ID NO: 4, 8 or 31 bind to the complement of SEQ ID NO: 1, for the sake of clarity, the target sites of the probe and forward primer are illustrated with reference to SEQ ID NO: 1.

Thus, the first shaded region represents the target site to which the forward primer SEQ ID NO: 4, 8 or 31 binds, ie. nucleotide residues –11 to 18 of the complement of SEQ ID NO: 1. The second shaded region represents the target site for the probe SEQ ID NO: 3, ie. nucleotide residues 66-82 within the complement of SEQ ID NO: 1. The third shaded region represents the target site to which the reverse primer SEQ ID NO: 5 binds, ie. nucleotide residues 214 to 240 of SEQ ID NO: 1.

Thus, Figure 1 illustrates that using a forward primer of SEQ ID NO: 4, 8 or 31, and a reverse primer of SEQ ID NO: 5, the resulting amplicon is 251 nucleotides long (spanning nucleotide residue -11 to nucleotide residue 240).

Figure 2 illustrates the relationship between the Mg219 gene (SEQ ID NO: 1), forward primers SEQ ID NOs: 9-19 and reverse primers SEQ ID NOs: 20-30.

In more detail, Figure 2 illustrates the 447 nucleotide sequence of Mg219 – ie. SEQ ID NO: 1 (central shaded bar, annotated every 100 nucleotides) plus 100 upstream nucleotides and over 500 downstream nucleotides. Shaded arrows 9 and 20 represent the target/ binding sites for the upstream and downstream sequencing primers SEQ ID NOs: 9 and 20, respectively. Shaded arrows 10-19 represent the target/ binding sites for the forward primers SEQ ID NOs: 10-

19. Shaded arrows 21-30 represent the target/ binding sites for the reverse primers SEQ ID NOs: 21-30. In this regard, although the forward primers bind to the complement of SEQ ID NO: 1, for the sake of clarity, the target sites of the forward primer are illustrated with reference to SEQ ID NO: 1.

Thus, it is evident from Figure 2 that primer SEQ ID NOs: 9 and 10 bind to target sequences located entirely within nucleotide sequences that are upstream of the complement of the Mg219 nucleic acid sequence. Likewise, primer sequences 20 and 21 bind to target sequences located entirely within nucleotide sequences that are downstream of the Mg219 nucleic acid sequence. In contrast, primer SEQ ID NOs: 11-18 bind to target sequences located entirely within the complement of the Mg219 nucleic acid sequence, and primer SEQ ID NOs: 22-29 bind to target sequences located entirely within the Mg219 nucleic acid sequence. However, the target site for primer SEQ ID NO: 19 overlaps the complement of the Mg219 nucleic acid sequence and (3) nucleotides located downstream of the complement of the Mg219 nucleic acid sequence. Likewise, the target site for primer SEQ ID NO: 30 overlaps the Mg219 nucleic acid sequence and (3) nucleotides located upstream of the complement of the Mg219 nucleic acid sequence.

Figure 3 illustrates the results of gradient PCR to determine optimal annealing temperature using primer sequences SEQ ID NO: 9 and SEQ ID NO: 20. The temperatures tested are as follows: A = 57.0°C, B = 57.3°C, C = 58.0°C, D = 58.8°C, E = 60.1°C, F = 61.8°C, G = 63.6°C, H = 65.5°C, I = 66.4°C, J = 67.2°C, K = control.

Figure 4 illustrates detection of the entire Mg219 gene using a span of oligonucleotide primers (forward primer SEQ ID NO: 9 with reverse primers selected from SEQ ID NOs: 20-30) across the full gene sequence. The primer pairs employed are as follows:

A = SEQ ID NO: 9 plus SEQ ID NO: 30
B = SEQ ID NO: 9 plus SEQ ID NO: 29
C = SEQ ID NO: 9 plus SEQ ID NO: 28

D = SEQ ID NO: 9 plus SEQ ID NO: 27
E = SEQ ID NO: 9 plus SEQ ID NO: 26
F = SEQ ID NO: 9 plus SEQ ID NO: 25
G = SEQ ID NO: 9 plus SEQ ID NO: 24
H = SEQ ID NO: 9 plus SEQ ID NO: 23
I = SEQ ID NO: 9 plus SEQ ID NO: 22
J = SEQ ID NO: 9 plus SEQ ID NO: 21
K = SEQ ID NO: 9 plus SEQ ID NO: 20

Figure 5 illustrates detection of the entire Mg219 gene using a span of oligonucleotide primers (reverse primer SEQ ID NO: 20 with forward primers selected from SEQ ID NOs: 9-19) across the full gene sequence. The primer pairs employed are as follows:

A = SEQ ID NO: 20 plus SEQ ID NO: 10
B = SEQ ID NO: 20 plus SEQ ID NO: 11
C = SEQ ID NO: 20 plus SEQ ID NO: 12
D = SEQ ID NO: 20 plus SEQ ID NO: 13
E = SEQ ID NO: 20 plus SEQ ID NO: 14
F = SEQ ID NO: 20 plus SEQ ID NO: 15
G = SEQ ID NO: 20 plus SEQ ID NO: 16
H = SEQ ID NO: 20 plus SEQ ID NO: 17
I = SEQ ID NO: 20 plus SEQ ID NO: 18
J = SEQ ID NO: 20 plus SEQ ID NO: 19
K = SEQ ID NO: 20 plus SEQ ID NO: 9
L = Negative control

Figure 6 illustrates amplification of Mg219 from 12 cultured *M. genitalium* strains (identified as 1-12, see Example 5 for key), using the primer pair SEQ ID NO: 9 and SEQ ID NO: 20.

Figure 7 presents an alignment of the Mg219 gene (plus upstream and downstream nucleotides) from *M. genitalium* strain G37 plus 16 cultured isolates, illustrating the binding sites for the forward and reverse sequencing

primers (SEQ ID NOs: 9 and 20), the forward assay primer (SEQ ID NO: 4 , 8 or 31), the reverse assay primer (SEQ ID NO: 5) and the probe (SEQ ID NO: 3). Each of the aligned sequences has been allocated a SEQ ID NO (see Example 5, below).

Figure 8 is a ClustalW (v1.4) multiple sequence alignment matrix, illustrating the sequence identity level between the Mg219 sequences (plus up- and down-stream nucleotides) illustrated in Figure 7. Alignment Score = 477772, Gaps Inserted = 30, Conserved Identities = 549. Pairwise Alignment Mode: Fast. Pairwise Alignment Parameters: ktup = 1, Gap Penalty = 3, Top Diagonals = 5 Window Size = 5. Multiple Alignment Parameters: Open Gap Penalty = 10.0, Extend Gap Penalty = 0.1, Delay Divergent = 40% Transitions: Weighted.

EXAMPLES

Example 1 – Standard Operating Procedure for detecting *M. genitalium* Mg219 gene

This Example details the procedures for the detection of *M. genitalium* Mg219 gene DNA from clinical material or cultures by Polymerase Chain Reaction (PCR) using fluorescent probes on the Corbett Rotor-gene real-time PCR machine.

Gloves and safety glasses were worn throughout, and all extraction procedures were carried out in a Class 1 Safety cabinet. Lower respiratory tract specimens such as sputa and lung biopsies were handled in a CL3 facility while other samples were processed in a Class 1 Safety Cabinet.

Preparing and running the PCR

Immediately prior to loading tubes into the Rotor-gene PCR machine, the template sample was added to the mixture in a designated PCR cabinet in the PCR clean room. PCR tubes containing mixtures prepared for cycling were capped before leaving the cabinet and conveyed to the Rotor-gene.

Materials

Sterile screw cap 1.5mL microfuge plastic tubes
Sputasol™. (Oxoid) freshly diluted in water (Sigma)
QIA-Amp DNA Mini Kit. (Qiagen cat no 51306)
Gilson pipettes: P1000, P200, P20, P10 (or equivalent)
Pre-sterile tips with filters for above (Rainin)
Corbett Rotagene Realtime PCR machine
0.5mL PCR tubes (flat top)
Tris-EDTA (TE) buffer 100 x concentrate (Sigma cat no T9285)
AmpliTaq Gold® DNA Polymerase, with GeneAMP (Applied Biosystems cat no 4311820)

AmpErase® Uracil N-glycosylase (UNG) (Applied Biosystems cat no N8080096)

Heating block 56±2°C.

Primers and probes: Adjusted to give 100μM stock solutions, aliquoted (50μL) and stored at -30°C or below.

M. genitalium NCTC 10195^{T} (Positive control) cloned stocks of DNA template (pGEMT- EASY genit219 POS 1, 2, 3, 4 or 5 in E. coli TOP10^{F1})

GeneQuant.II (Pharmacia Biotech)

Nuclease-free water (Promega cat no P119C)

Herring sperm DNA

dNTP Blend, 12.5mM with dUTP (Applied Biosystems cat no N8080270)

Procedures

Samples received for PCR may be tested by culture for Mycoplasmas and Ureaplasmas (culture samples prior to commencing DNA extraction).

DNA was extracted using a conventional DNA extraction protocol (depending on the circumstances eg. nature and volume of sample, urgency of request etc. different DNA extraction protocols may be used).

DNA from urine was extracted via centrifugation at 20,000 g for 15 mins, and then the pellet was re-suspended in 0.4mL PBS and heated at 100°C for 5 mins (alternatively, the magnapure compact may be used directly on 0.5mL heat inactivated urine (95°C±10°C) without pre-centrifugation). Swabs were extracted by agitating in 0.4mL PBS and following the same protocol as above.

PBS or water was extracted in tandem with the clinical sample and included in all further tests to ensure cross contamination did not occur.

A 1:10 dilution of the extracted DNA was prepared in nuclease free water. For unusual and highly cellular samples such as tissue a 1:10 and 1:100 dilution was prepared.

Preparation of cloned *M. genitalium* standard for regression calculation

Because of the problems associated with growing sufficient *Mycoplasma* to generate stocks of standard DNA, a cloned template DNA sequence was prepared.

The concentration and purity of the plasmid solution in TE buffer was estimated using GeneQuant (against TE blank).

The plasmid solution was adjusted to known concentration (100ng/mL) in Herring sperm DNA at 10ng/μL (in TE buffer pH 8.0), aliquoted (25μL), and stored at –20±10°C.

Immediately prior to use, serial dilutions were made (in PCR quality water) to cover the expected dynamic range of the clinical specimens and a water blank. 5ng, 0.5ng, 0.05ng and 0.05ng serial dilutions were used in each run.

Preparation of primers and probes

The primers used in this assay yield a positive product of 251 bp for *M. genitalium* and the probe binds from bp 66 to 82 inside the Mg219 gene.

Forward Primer = SEQ ID NO: 4, 8 or 31; Reverse Primer = SEQ ID NO: 5.

Primer aliquots were re-suspended in TE to give 100pMμL^{-1}, aliquoted, assigned batch numbers and stored at –20±10°C or below.

Probes (SEQ ID NO: 3, [at the 3 nmol scale]) were re-suspended in TE to give 100pMμL^{-1}, aliquoted, assigned batch numbers and stored at –20±10°C.

The probes are light sensitive and should therefore be handled in dark/amber microfuge tubes at all times.

The working primer/ probe mix contains 2.5pmolμL^{-1} (2.5μM) of each primer, giving a final concentration of 0.5μM each and 1 pmol μL^{-1} (10μM) of *M. genitalium* probe, giving a final concentration of 0.2μM. To make the mix, the following components were added:

µL	(4mL total)
3760	H_20
100	Primer Mg219F (SEQ ID NO: 4, 8 or 31)
100	Primer Mg219R (SEQ ID NO: 5)
40	Probe Mg219 (SEQ ID NO: 3)

The primer/ probe mix was distributed in 1000μL aliquots in black tubes, assigned batch numbers and an expiry date of 12 months and stored at -20±10° C.

Preparation of PCR reagent mixture.

A stock mastermix was made for 1000 reactions and aliquoted giving enough mastermix for a full carousel (37 reactions). The stock mastermix was stored at -20±10°C and given expiry date of 12 months from formation. The reagents were mixed gently by pipetting before use.

Mastermix	Per sample	Master mix	Final conc.
H_2O	3.55µL	3.55mL	
$MgCl_2$ 25 mM	4µL	4mL	5µM
dNTPs 12.5mM	1µL	1mL	0.625mM
Primer/ Probe mix	4µL	4mL	
	(0.2µM MG219 probe, 0.5µM primer)		
Amplitaq gold 10 x buffer	2µL	2mL	x1

538.35μL were distributed into each tube and frozen at -20±10°C.

On day of use 1 vial was defrosted and the following were added to 37 reactions:

Amplitaq gold polymerase	0.25μL	9.25μL	1.25 unit
Uracil-DNA glycosylase	0.2μL	7.4μL	1 unit

Total	15μL

15μL was distributed into each tube.

All manipulations were performed at 2-8°C. A water blank was included in every PCR experiment. Water was pipetted after other solutions to control for cross contamination. The clinical sample, dilutions and extracted water control were included, pipetted and sealed before adding positive controls.

5μL of the relevant sample were added and the tubes were closed (giving a final volume of 20μL). A standard curve of positive control DNA was included. The PCR machine was loaded, making sure that the tubes were closed properly by firmly pressing down the lid and always using the locking rings provided to secure lids on both 0.1 and 0.2mL tubes.

NOTE: Calibration measures the fluorescence in the 1st tube for all channels and adjusts fluorescence respectively to take account of background. Therefore the first tube must contain all probes used in the assay. Baseline fluorescence should be ~20, for quenched FRET assays where a decrease in fluorescence is expected it should be 50-70. Set calibration to be taken at beginning of first cycle to relevant channel.

Detection of PCR Products

Setting up the Rotorgene PCR machine:

The PCR machine was set up, samples were put in and the programme was entered/ opened. Flat tubes and the required programme channels were set (Cy5 – channel 4) and the details of specimens and standards in respective carousel positions were entered. Save and 'start'. The entire carousel was filled with tubes, even if not included in the analysis. The assay was run.

Cycling conditions

45 cycles

	Temp	Time	Transition rate	Acquisition mode
Denaturation	95°C	5 min	20 secs	None
Quantification				
Segment 1	95°C	15 sec	20°C/sec	None
2	60°C	60 sec	20°C/sec	Single
Melting Curve				
Segment 1	95°C	0 sec	20°C/sec	None
2	45°C	2 sec	20°C/sec	None
3	80°C	0 sec	0.1°C/sec	Cont

Cooling

The Rotor-gene automatically cools

Data analysis

By clicking the 'analysis' button and double clicking on the Cy5 channel, the software gives the standard curve an efficiency value. If they are unexpectedly high, the assay should be repeated.

Example 2 – Validation of the assay

The purpose of the validation was to determine the following:

1) Specificity – other *Mycoplasma* species and bacterial species from same niche
2) Sensitivity – determine LDL
3) Known positives and negatives
4) Optimal extraction method
5) Alternative PCR platforms
6) Varying sample types
7) Reproducibility
8) Comparison to MgPa assay

1) Specificity

Species	NCTC/ strain identifier	DNA conc	Test result
Commensals			
Lactobacillus vaginalis	12197	-	Neg
Prevotella (Bacteroides) bivius	11156	1µgmL^{-1}	Neg
Mobiluncus curtisii	11657	-	Neg
Mobiluncus mulieris	11658	1µgmL^{-1}	Neg
Gp B *Streptococcus (S. pyogenes)*	12067	-	Neg
Gp A *Streptococcus (S. algalactiae)*	12906	-	Neg
Echerishia. coli	9001	1µgmL^{-1}	Neg
Staphylococcus	8532	-	Neg
Pathogens			
Neisseria gonorrhoea	H060160180 Clinical sample	-	Neg
Chlamydia trachomatis	L2, Cell culture DNA	-	Neg
Treponema palidum	H05510345 Clinical sample	-	Neg
Treponema palidum	Nichols or Newmarket cell culture strain	-	Neg
Heamophilus ducreyi	PCR control	-	Neg
HSV1	PCR control	-	Neg
Mollicute species			
A. laidlawii		1µgmL^{-1}	Neg
M. hominis		1µgmL^{-1}	Neg
M. pneumoniae		1µgmL^{-1}	Neg
M. amphoriforme		1µgmL^{-1}	Neg
M. fermentans		1µgmL^{-1}	Neg
M faucium		1µgmL^{-1}	Neg
M. penetrans		1µgmL^{-1}	Neg
M. pirum		1µgmL^{-1}	Neg
M. spermatophilum		1µgmL^{-1}	Neg
M. primatum		1µgmL^{-1}	Neg
M. salivarium		1µgmL^{-1}	Neg
M. buccale		1µgmL^{-1}	Neg
M. orale		1µgmL^{-1}	Neg
M. lipophilum		1µgmL^{-1}	Neg
U. parvum (Sero 1)		1µgmL^{-1}	Neg
U. urealyticum (Sero 12)		1µgmL^{-1}	Neg

Summary – the assay does not amplify other bacteria or mollicutes

2) Sensitivity

Standard conc (μgmL^{-1})	Result	Cp (average)
100	Pos	15.28, 15.16 (15.22)
10	Pos	19.2, 19.07 (19.14)
1	Pos	24.37, 21.37 (24.37)
-1	Pos	27.33, 27.28 (27.3)
-2	Pos	31.27, 31.41, 31.08,31.12 (31.22)
-3	Pos	34.46, 35.33, 34.84, 36.27 (35.17)
-4	Pos	37.87, 39.17,38.18,- (38.41)
-5	Neg	38.56, -,-,-
-6	Neg	-,-,-,-
-7	Neg	-,-,-,-

LDL calculations:

The lowest detectable limit of detection was 5µL of a 10-4 dilution of a 1μgmL^{-1} extract of DNA. This corresponds to 0.0001μgmL^{-1} or 100pg, 0.5pg per 5 µL in reaction and 825 gene copies per reaction. The assay could detect DNA at one further dilution (at 82.5 gene copies per reaction), however reproducibility was then affected.

The assay LDL is = 0.5pg per reaction

The point at which a real-time PCR is determined as positive can be known as a crossing point (Cp). A lower Cp indicates a larger amount of target DNA in the sample or more efficient amplification. For all positive specimens to date the mean Cp was 34 in the Mg219 assay of the present invention, and 36 in the known MgPa assay.

This indicates that the Mg219 assay of the present invention may be more sensitive that the known MgPa assay.

3) Known positives/ negatives

Sample type	Sample number/ID	Expected Result (Mg219 assay)	Actual result (MgPa assay)
Urethral isolate (J. Tully)	M30 10-4	P	P*
Urethral isolate (J. Tully)	M 30 early 10-2	P	P*
Human throat (J. Tully)	TW10-5G 10-4	P	P*
Human throat (J. Tully)	TW10-6 G10-4	P	P*
Human Throat (J. Tully)	R-32G 10-3	P	P*
Human throat (J. Tully)	TW 48-5G 10-4	P	P*
Urethra (J. Jensen)	M 2300 10-4	P	P*
Urethra (J. Jensen)	M2321 10-3	P	P*
Urethra (J. Jensen)	M2341 10-3	P	P*
Urethral isolate (D.T. Robinson)	G37 10-4	P	P*
Urethra (J. Jensen)	M 2288 10-3	P	P*
Synovial fluid (J. Tully)	UTMB-10G 10-4	P	P*
Water	N/A	N	N
PBS	N/A	N	N

P = positive in Jensen *et al.* (2004) publication

Summary – all known positives are detected.
Known negatives are not reactive in the assay

4) Optimal extraction method

Extraction of *M. hominis* spiked urine was performed using varying methods (boiling, qiagen, magnapure compact). The magnapure compact gave the optimal results.

- Samples are best stored at 4°C, not at room temperature.
- DNA was detectable for up to 28 days.

- The addition of RND/DNA protect enhances detection and eliminates the temperature differences - but does not alter results and is not therefore cost effective.
- Concentration of sample prior to extraction of DNA is recommended.

5) Alternative PCR platforms

The Mg219 assay has been tested on the Roche Lightcycler and the ABI Taqman 7700. The assay performed satisfactorily on both platforms (positive controls were amplified and detected whilst negative controls were not).

These results indicate that the chemistry used in the Mg219 assay is transferable to other real-time platforms.

6) Varying sample types (comparison with MgPa assay)

Sample	**Mg219 (actual result)**	**MgPa assay (expected result)**
Urine	0/112 positive	0/112 positive (1 FP)
Rectal swabs	2/22 positive	2/22 positive
Anal swab	0/1 positive	0/1 positive
Urethral swabs	0/13 positive	0/13 positive
Genital ulcer swabs	0/18 positive	0/18 positive
Penile swab	1/1 positive	0/1 positive
Liquid based cytology samples – cervical swab	0/30 positive	0/30 positive
Vaginal swabs	0/35 positive	0/35 positive
Respiratory samples	0/38 positive	0/38 positive (1 FP)
Endocervical samples	5/309 positive	5/309 positive (7 FP)
Defined samples total	**8/579 positive (1.38%)**	**7/579 positive (1.21%) (9 FP)**

FP – false positive – specimens positive only in the MgPa assay are thought to be false positive results (see below).

7) Reproducibility

Mg219 controls, CT values when run on several days

Date	Expt number	1/10	1/100	1/1000
16/2/6	0236	13.86	17.17	22.8
17/2/6	0240	12.39	16.89	22.23
24/2/6	0247	11.24	16.63	22.14
02/3/6	0249	10.99	17.21	24.01
03/3/6	0251	12.83	19.08	23.84
7/4/6	0255	11.66	17.44	23.23
11/5/6	0258	12.84	15.81	19.31
	AVERAGE	**12.26**	**17.18**	**22.51**
	STDEV	1.02	1.00	1.59
	Boundaries (+/- 2 x STDEV)	10.22-14.30	15.19-19.17	19.34-25.68

Note all runs are within average +/- 2 x STDEV.

8) Comparison to known MgPa assay (Jensen *et al.*, 2004)

	Known MgPa assay	New HPA MG219 assay
Specificity	Does not cross react with any human Mycoplasma species	Does not cross react with any human Mycoplasma species
	Does not cross react with bacteria and viruses listed above	Does not cross react with bacteria and viruses listed above
		Gene and primers/ probes do not align to any other genes with significant homology on search of BLAST and other public databases
		Using ORFANAGE program MG219 is found only in *M. genitalium* and not in other published whole genome sequences
		Using MOLLIGEN whole genome comparison, MG219 is not found in any other Mycoplasma species.
Sensitivity	Lower detection limit = 0.5pg per reaction of *M. genitalium* DNA.	Lower detection limit = 0.5pg per reaction of *M. genitalium* DNA (same as MgPa assay).
Known positives	Detects all known positives tested (culture positive)	Detects all known positives tested (culture positive)
Known negatives	Does not amplify known negatives	Does not amplify known negatives
Sample	**MgPa assay (expected result)**	**Mg219 (actual result)**
Urine	0/112 pos (1FP)	0/112 pos
Rectal swabs	2/22 pos	2/22 pos
Anal swab	0/1 pos	0/1 pos
Urethral swabs	0/13 pos	0/13 pos
Genital ulcer swabs	0/18 pos	0/18 pos
Penile swab	0/1 pos	1/1 pos (1FP)
LBC samples – cervical swab	0/30 pos	0/30 pos
Vaginal swabs	0/35 pos	0/35 pos
Respiratory samples	0/38 pos (1FP)	0/38 pos
Endocervical samples	5/309 pos (7FP)	5/309 pos
Defined samples total	**7/579 pos (9FP)**	**8/579 pos (1FP)**
Undefined samples (query type)	**4/155 pos (11FP)**	**4/155 (0FP)**

Example 3 – Epidemiological data (UK population)

To date, 548 specimens have been tested, of which 9 (1.65%) were positive in the MG219 assay (see Table below). This indicates a level of approximately 1.65% infected individuals within the UK population.

Epidemiological information from research study specimens		
	Mg219 (actual result)	**Known MgPa assay (expected result)**
Cornwall NHS Trust	0/134 pos (0%)	0/134 pos (3 FP)
University Hospital College London	5/274 pos (1.82%)	5/274 (5 FP)
Northampton General Hospital	4/140 pos (2.86%)	4/140 pos (6 FP)
Research study specimens total	**9/548 pos (1.65%)**	**9/548 pos (14 FP)**

The samples are taken from both symptomatic and asymptomatic patients and therefore the true level of infection in specific clinical groups (such as those without symptoms or with urethritis) may be lower or higher than reflected herein. Also, these samples are in the main from 3 geographical areas (Cornwall, London, Northampton) with levels varying from 0%, 1.82% to 2.86%. Larger scale screening will indicate if geographical differences occur in infection levels.

False positive specimens in the known MgPa assay:

To date we have 20 specimens which are positive according to the MgPa assay described by Jensen *et al.* (2004) and negative in the Mg219 assay of the present invention. On repeat testing only 2 of the 20 were positive in the MgPa test, which indicates that they are likely to be false positives in the Jensen *et al.* MgPa test. In this regard, the present inventors have been unable to repeat the MgPa positive result and have not had any positives using two other PCR methods (Cadieux N. *et al.* (1993) J Gen Microbiol. 1993

Oct;139(10): 2431-7, and MG192 PCR by Musatovova O. *et al.*, (2006) J Clin Microbiol. 44(2):598-603).

The mean Cp for these samples was 39, which could be indicative of inefficient amplification due to either lower amounts of target DNA or non-specific amplification. The former is unlikely, because all samples positive in both assays had lower mean Cp in the MG219 assay of the present invention than in the known MgPa assay.

This indicates that the Mg219 assay of the present invention is more sensitive than the known MgPa assay for the detection of *M. genitalium*.

Example 4 – Further primer design

A. Designing primers for PCR of whole of MG219

Two primers were designed binding upstream and downstream of the Mg219 nucleic acid sequence. As illustrated in the Table below, the up-stream, forward primer SEQ ID NO: 9 mapped to position -76 to -54 (upstream of the start codon) and the downstream, reverse primer SEQ ID NO: 20 mapped to position 503 to 525 (downstream of the stop codon) of MG219. The annealing temperature of the primers was calculated using the formula: Tm=4*(number of G's and C's) + 2*(number of A's and T's).

Primer	Sequence	Position
SEQ ID NO: 9	CCACTTAACTTTATTACCCGTCC	-76 to -54
SEQ ID NO: 20	GATTAACCCCAGGTAGTTCTTCC	+503 – 525

To confirm the suitability of the primers, BLAST analysis was carried out for each of the primers against the *M. genitalium* G37 genome. Selecting the "Advanced Blast" parameters, small primer sequences showed unique binding

sites and enabled determination of homologous sequences elsewhere on the genome, and prediction of possible multiple amplification products upon PCR.

B. Determining Optimum annealing temperature for PCR

To optimise the annealing temperature and avoid multiple PCR products, gradient PCR was performed using the primers SEQ ID NO: 9 and SEQ ID NO: 20, ranging from 57-68°C. This was performed on a MJ Research DNA engine, which does not perform a linear gradient across the block (Figure 3).

No other products were observed – indicating that only the MG219 gene was amplified. The resulting amplicons were used to determine the sequence of the MG219 gene and to amplify across the gene.

C. Detection of Mg219

Forward primers SEQ ID NOs: 10-19 were designed starting from position -53 relative to the start codon and mapping consecutively to 3 bases beyond the stop codon, to cover the entire gene sequence of Mg219 (see Table below).

Forward Primer SEQ ID NOs:	Sequence	Position
10	TGTTTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT	-53 to -1
11	ATGCGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA	1 – 50
12	TCTAAAACTCCAAAAGGTGTGGATCGAGCGGCATGTTGATCAAGATGAAC	51 – 100
13	TTAGTTTAACAACTACTGCAGTTGAACTTAAAAAGAGTGATGAACAAAAA	101 – 150
14	CCTGTTGCCATTAAAAGTAGTGACTTTATTGGTCATGAAGAGTTAATCTC	151 – 200
15	TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC	201 – 250
16	CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA	251 – 300
17	ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC	301 – 350
18	TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC	351 – 400
19	CAACCCAAATTCAAACAAAAAAAAGCAATAAAAAAACCAGATCTTAATCT	401 – +450

Reverse primer SEQ ID NOs: 21-30 were designed starting from position +502 from the start codon and they map consecutively upstream towards the start codon (see Table below).

Reverse Primer SEQ ID NOs:	Sequence	Position
21	TATTCCTTTCCAGTTTTTAGTTAAAACTACTGTTGTTAACACTAAAAAAACCAGA	+448 – +502
22	TTAAGATCTGGTTTTTTTATTGCTTTTTTTTGTTTGAATTTGGGTTGGTT	398 – 447
23	TAGCTTTGGTTTGTTGAACTTGTTTTGATTTGGGCTTGGATTGTTTAGGT	348 – 397
24	TTTGTTTGTTTAGTTGATTTACTAGTTGGTTTTGATTTAACAGGAGTTTT	298 – 347
25	CTTTTTAGTTGCTTTTGGTTTTGCTTTAACTGGAGGAATAACTGCTGGTT	248 – 297
26	GATCAATCTCTTTAACAACAGGGGTTGGGATTAGTAAAACTGGAACAGAG	198 – 247
27	ATTAACTCTTCATGACCAATAAAGTCACTACTTTTAATGGCAACAGGTTT	148 – 197
28	TTGTTCATCACTCTTTTTAAGTTCAACTGCAGTAGTTGTTAAACTAAGTT	98 – 147
29	CATCTTGATCAACATGCCGCTCGATCCACACCTTTTGGAGTTTTAGATCA	48 – 97
30	CTATCACTATTCATTATGGGTATTTTTTTCAAGTAACTGGTGCGCATAAT	-3 – 47

Individual PCRs were performed using forward primer SEQ ID NO: 9 plus any one of reverse primers SEQ ID NOs: 21-30; and also using forward primer SEQ ID NO: 20 plus any one of reverse primers SEQ ID NOs: 10-19.

As illustrated in Figures 4 and 5, PCR amplification using the primer pairs described above detected the whole Mg219 gene.

Example 5 – Conservation of the Mg219 gene

PCR amplification of the entire Mg219 gene and flanking regions (upstream and downstream sequences) was performed using primers SEQ ID NO: 9 and SEQ ID NO: 20 on twelve cultured *M. genitalium* isolates (see Table below) and the type strain MG37 using Roche High Fidelity Taq.

The resultant amplified fragments were all of the same size (except isolates 9 and 11, which appeared slightly larger on agarose gel electrophoresis) – see Figure 6.

Cultured Isolate Number	*M. genitalium* strain name	Origin
1	M30 10-4	Urethra
2	M30 early 10-2	Urethra
3	TW10-5G 10-4	Throat
4	TW10-6G 10-4	Throat
5	R-32G 10-3	Throat
6	TW48-5G 10-4	Throat
7	M2300 10-4	Urethra
8	M2321 10-3	Urethra
9	M2341 10-3	Urethra
10	G37 10-4	Urethra
11	M2288 10-3	Urethra
12	UTMB 10G 10-4	Synovial fluid
G37	NCTC type strain	Urethra

The resulting amplified fragments were purified and the concentrations determined. Both strands of the PCR products were sequenced using the oligonucleotide primers SEQ ID NO: 9 and SEQ ID NO: 20 in triplicate. Sequencing was performed using the Becton-Dickinson sequencing kit following the manufacture's instructions and using the CEQ8000 DNA sequencer instrument and the consensus sequence for each strain was determined using MapVector and ClustalW software.

Sequences were aligned and the alignment showed high conservation of the MG219 gene (see Table below and Figures 7 and 8). Only 40 bases differed from the 569 bases of the control G37 strain and these differences were found only within strains 9 and 11. Of these, 30 bases consisted of an insertion in isolates 9 and 11 between residues 504 and 505 of the control strain.

Further subsequent sequence analysis of four positive samples (identified as +ve MG 2, +ve MG 3, +ve MG 5 and +ve MG 6, below) has determined that positive clinical specimens are also highly conserved. The insertion sequence is also present in the sequence from Sample +ve MG 6.

Isolate Number	SEQ ID NO:	MG219 identity to G37 strain
1	32	100%
2	33	100%
3	34	100%
4	35	100%
5	36	100%
6	37	100%
7	38	100%
8	39	100%
9	40	93.1% (40/569)
10	41	100%
11	42	93.2% (39/569)
12	43	100%
+ve MG 2	44	99.8% (1/569)
+ve MG 3	45	100%
+ve MG 5	46	100%
+ve MG 6	47	93.2% (39/569)
G37	48	N/A

2007266857 19 Nov 2012

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method for detecting *M. genitalium* nucleic acid in a sample, comprising:

(i) amplifying a nucleic acid sequence comprising a fragment of SEQ ID NO: 1, wherein said fragment of SEQ ID NO: 1 is at least 25 consecutive nucleotides of SEQ ID NO: 1; and

(ii) detecting said amplified nucleic acid sequence.

2. A method according to Claim 1, wherein said amplifying step comprises contacting said sample with forward and reverse oligonucleotide primers, wherein said primers bind to target sites in the *M. genitalium* nucleic acid, under conditions suitable to promote amplification of said nucleic acid sequence comprising said fragment of SEQ ID NO: 1.

3. A method according to Claim 2, wherein said forward and reverse oligonucleotide primers are at least 5 nucleotides long.

4. A method according to Claim 3, wherein said forward and reverse oligonucleotide primers are no more than 50 nucleotides long.

5. A method according to any one of Claims 1 to 4, wherein said detecting step comprises:

(i) contacting said amplified nucleic acid sequence with a probe, wherein the probe binds to a target site within said amplified nucleic acid sequence, or the complement thereof; and

(ii) detecting binding of said probe to said amplified nucleic acid sequence.

6. A method according to Claim 5, wherein said probe target site is located:

(a) between nucleotide residues 50 and 100 of SEQ ID NO: 1 or the complement thereof;

2007266857 19 Nov 2012

(b) between nucleotide residue 1 and nucleotide residue 447, 440, 420, 400, 350, 300, 250, 200, 150, 100 or 50 of SEQ ID NO: 1 or the complement thereof; or

(c) between nucleotide residue 10, 15, 25, 50, 75, 100, 150, 200, 250, 300, 350 or 400 of SEQ ID NO: 1 or the complement thereof and nucleotide residue 447 of SEQ ID NO: 1 or the complement thereof;

(d) between nucleotide residue 10, 15 or 25 and nucleotide 50, 100 or 150 of SEQ ID NO: 1 or the complement thereof;

(e) between nucleotide residue 50, 75 or 100 and nucleotide 150, 200 or 250 of SEQ ID NO: 1 or the complement thereof;

(f) between nucleotide residue 150 or 200 and nucleotide 300 or 350 of SEQ ID NO: 1 or the complement thereof;

(g) between nucleotide 250 or 300 and nucleotide 350 or 400 of SEQ ID NO: 1 or the complement thereof.

7. A method for detecting *M. genitalium* nucleic acid in a sample, comprising:

(i) contacting said sample with an oligonucleotide probe, wherein the probe binds to a target site within SEQ ID NO: 1, or the complement thereof; and

(ii) detecting binding of said probe to said target site;

wherein said target site is located:

(a) between nucleotide residues 50 and 100 of SEQ ID NO: 1 or the complement thereof;

2007266857 19 Nov 2012

(b) between nucleotide residue 1 and nucleotide residue 447, 440, 420, 400, 350, 300, 250, 200, 150, 100 or 50 of SEQ ID NO: 1 or the complement thereof; or

(c) between nucleotide residue 10, 15, 25, 50, 75, 100, 150, 200, 250, 300, 350 or 400 of SEQ ID NO: 1 or the complement thereof and nucleotide residue 447 of SEQ ID NO: 1 or the complement thereof;

(d) between nucleotide residue 10, 15 or 25 and nucleotide 50, 100 or 150 of SEQ ID NO: 1 or the complement thereof;

(e) between nucleotide residue 50, 75 or 100 and nucleotide 150, 200 or 250 of SEQ ID NO: 1 or the complement thereof;

(f) between nucleotide residue 150 or 200 and nucleotide 300 or 350 of SEQ ID NO: 1 or the complement thereof;

(g) between nucleotide 250 or 300 and nucleotide 350 or 400 of SEQ ID NO: 1 or the complement thereof.

8. A method according to Claim 6 or 7, wherein the probe binds to a target site from nucleotide residue 50 to 100 of the nucleic acid strand complementary to SEQ ID NO: 1.

9. A method according to Claim 6 or 7, wherein the probe binds to a target site from nucleotide residue 60 to 85 of the nucleic acid strand complementary to SEQ ID NO: 1; or wherein the probe binds to a target site from nucleotide residue 66 to 82 of the nucleic acid strand complementary to SEQ ID NO: 1.

10. A method according to any one of Claims 5 to 9 wherein said probe is at least 15 nucleotides long and up to 50 nucleotides long.

11. A method according to Claim 10 wherein the probe comprises the nucleic acid sequence of SEQ ID NO: 3, or a fragment thereof having at least 15 nucleotides, or a variant thereof having at least 75% sequence identity thereto.

2007266857 19 Nov 2012

12. A method according to Claim 10 wherein the probe consists of the nucleic acid sequence of SEQ ID NO: 3, or a fragment thereof having at least 15 nucleotides, or a variant thereof having at least 75% sequence identity thereto.

13. A method according to Claim 11 wherein the probe comprises the nucleic acid sequence of SEQ ID NO: 3.

14. A method according to Claim 12 wherein the probe consists of the nucleic acid sequence of SEQ ID NO: 3.

15. An *in vitro* method for quantitating *M. genitalium* pathogen load in a sample of interest, comprising:

(a) carrying out a method according to any of Claims 1 to 14 on said sample of interest; and

(b) carrying out a method according to any of Claims 1 to 14 on a test sample of predetermined known *M. genitalium* pathogen load; and

(c) comparing the signal detected from the sample of interest with the signal detected from the test sample;

and thereby quantitating *M. genitalium* pathogen load in the sample of interest.

16. An *in vitro* method for determining drug efficacy over the course of a period of drug therapy, comprising:

(a) carrying out a method according to any of Claims 1 to 14 on a first sample obtained at a first time point within or prior to the period of drug therapy;

(b) carrying out a method according to any of Claims 1 to 14 on a sample at one or more later time points within the period of drug therapy; and

2007266857 19 Nov 2012

(c) comparing the signal detected from the first sample with the signal detected from the one or more later samples;

and thereby determining drug efficacy over the course of the period of drug therapy.

17. An *in vitro* method for detecting and/ or monitoring the development of resistance of *M. genitalium* to a drug, comprising:

(a) carrying out a method according to any of Claims 1 to 14 on a first sample obtained at a first time point within or prior to exposure of the *M. genitalium* to the drug;

(b) carrying out a method according to any of Claims 1 to 14 on a sample at one or more later time points following exposure of the *M. genitalium* to the drug; and

(c) comparing the signal detected from the first sample with the signal detected from the one or more later samples;

wherein the absence of a reduction in the signal, or an increase in the signal, indicates that the *M. genitalium* have developed resistance to the drug.

18. An oligonucleotide probe for use in a method according to any one of claims 1 to 17, wherein:

(a) said probe comprises the nucleic acid sequence of SEQ ID NO: 3, or

(b) said probe consists of the nucleic acid sequence of SEQ ID NO: 3, or a fragment thereof having at least 15 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto.

19. A forward primer for use in a method according to any one of Claims 1 to 17, wherein said forward primer comprises a nucleic acid sequence selected from:

2007266857 19 Nov 2012

(a) SEQ ID NOs: 4, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19 or 31, or a fragment thereof, having at least 20 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto; or

(b) SEQ ID NOs: 13, or a fragment thereof having at least 30 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto.

20. A forward primer according to Claim 19, wherein said forward primer comprises a nucleic acid sequence selected from:

(a) SEQ ID NOs: 4, 8, or 31, or a fragment thereof having at least 20 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto;

(b) SEQ ID NO: 10, or a fragment thereof having at least 40 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto.

21. A forward primer according to Claim 20, wherein said forward primer comprises a nucleic acid sequence selected from SEQ ID NOs: 4, 8 or 31.

22. A reverse primer for use in a method according to any of Claims 1 to 17, wherein said reverse primer comprises a nucleic acid sequence selected from:

(a) SEQ ID NOs: 5, 20, 21, 22, 23, 24, 26, 27, 28, 29 or 30, or a fragment thereof having at least 20 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto; or

(b) SEQ ID NO: 25, or a fragment thereof having at least 30 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto.

23. A reverse primer according to Claim 22, wherein said reverse primer comprises a nucleic acid sequence selected from:

2007266857 19 Nov 2012

(a) SEQ ID NO: 5, or a fragment thereof having at least 20 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto; or

(b) SEQ ID NOs: 26 or 28, or a fragment thereof having at least 40 consecutive nucleotides thereof, or a variant thereof having at least 90% sequence identity thereto.

24. A reverse primer according to Claim 23, wherein said reverse primer comprises the nucleic acid sequence SEQ ID NO: 5.

25. A pair of oligonucleotide primers, comprising a forward primer according to any one of Claims 19 to 21 and a reverse primer according to any one of Claims 22 to 24.

26. A kit for detection of *M. genitalium* nucleic acid comprising a probe according to Claim 18, and further comprising a forward primer according to any one of Claims 19 to 21 and/ or a reverse primer according to any one of Claims 22 to 24.

27. A method according to Claim 1, substantially as hereinbefore described.

28. An *in vitro* method according to any one of Claims 15 to 17, substantially as hereinbefore described.

29. A probe according to Claim 18, substantially as hereinbefore described.

30. A primer according to any one of Claims 19 to 24, substantially as hereinbefore described.

31. A kit according to Claim 26, substantially as hereinbefore described.

Figure 1

-11 catagttcatt 1 atgcgcacca gttacttgaa aaaaataccc ataatgaata gtgatagtga tctaaaactc
61 caaaaggtgt ggatcgagcg gcatgttgat caagatgaac ttagtttaac aactactgca
121 gttgaactta aaaagagtga tgaacaaaaa cctgttgcca ttaaaagtag tgactttatt
181 ggtcatgaag agttaatctc tgttccagtt ttactaatcc caacccctgt tgttaaagag
241 attgatcaac cagcagttat tcctccagtt aaagcaaaac caaaagcaac taaaaagaaa
301 actcctgtta aatcaaaacc aactagtaaa tcaactaaac aaacaaaacc taaacaatcc
361 aagcccaaat caaaacaagt tcaacaaacc aaagctaaac caacccaaat tcaaacaaaa
421 aaaagcaata aaaaaaccag atcttaa.

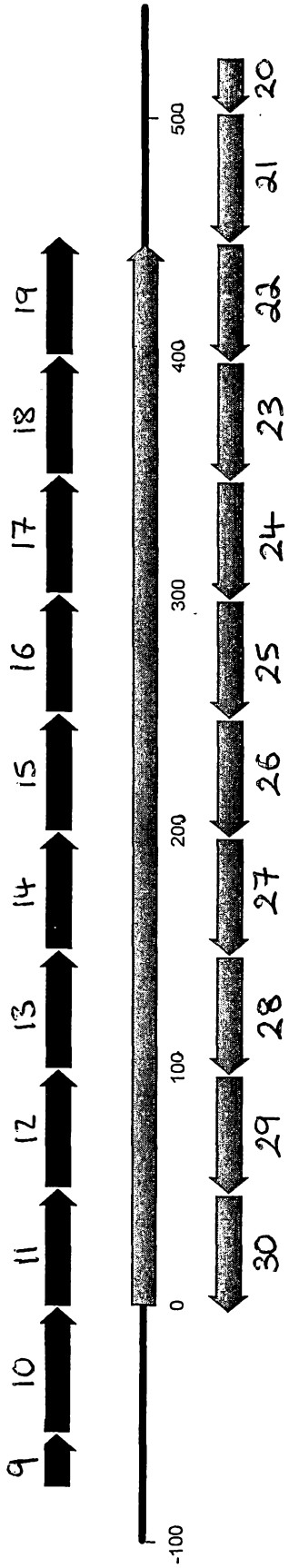

-100
0
100
200
300
400
500
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

Figure 2

Figure 3

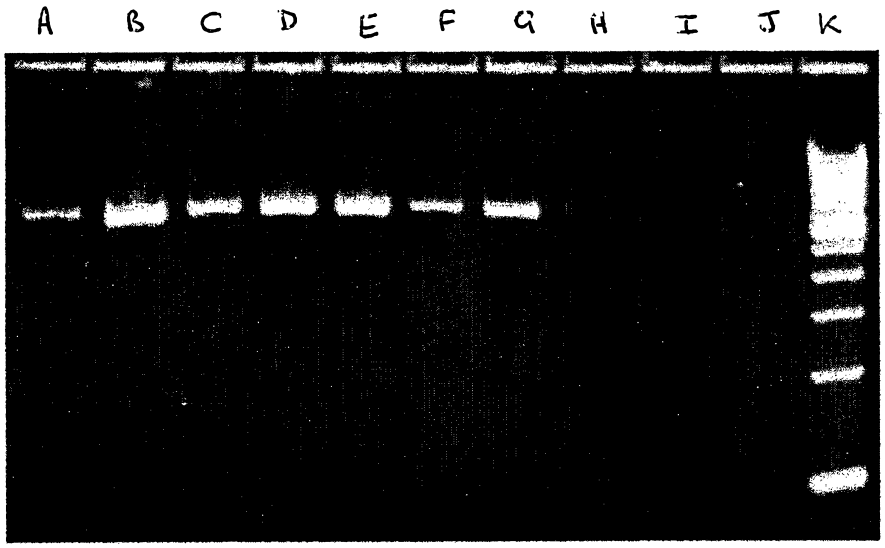


Figure 4

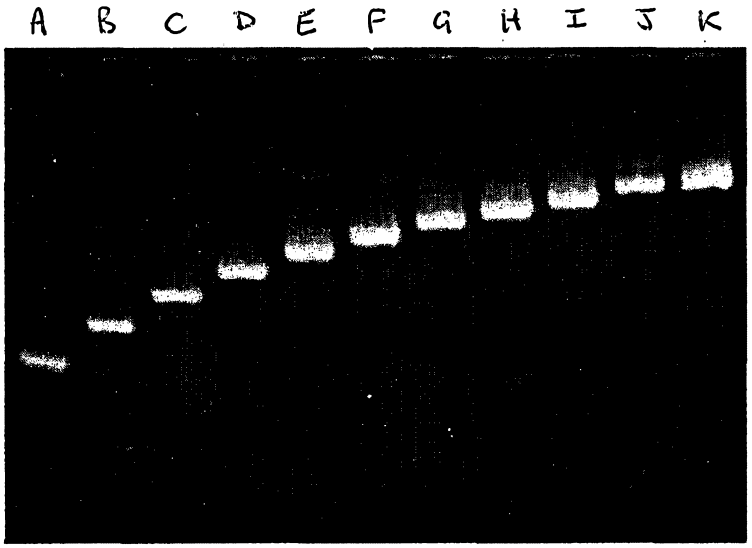


Figure 5

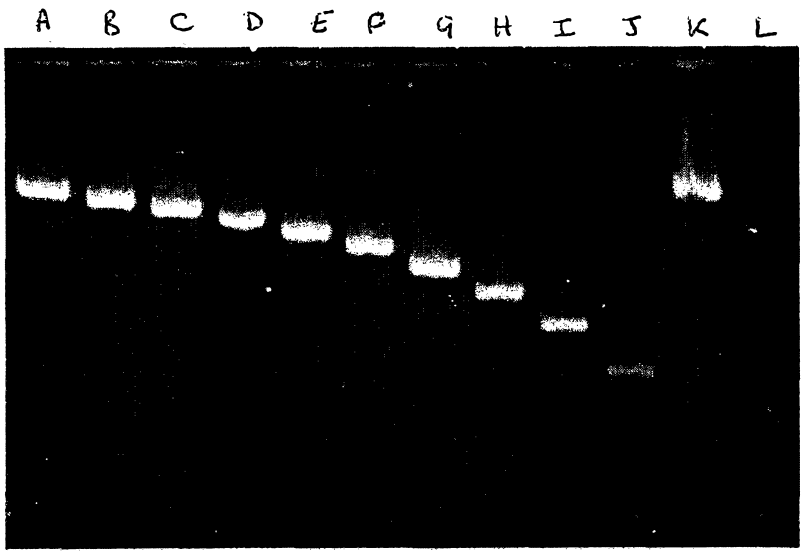
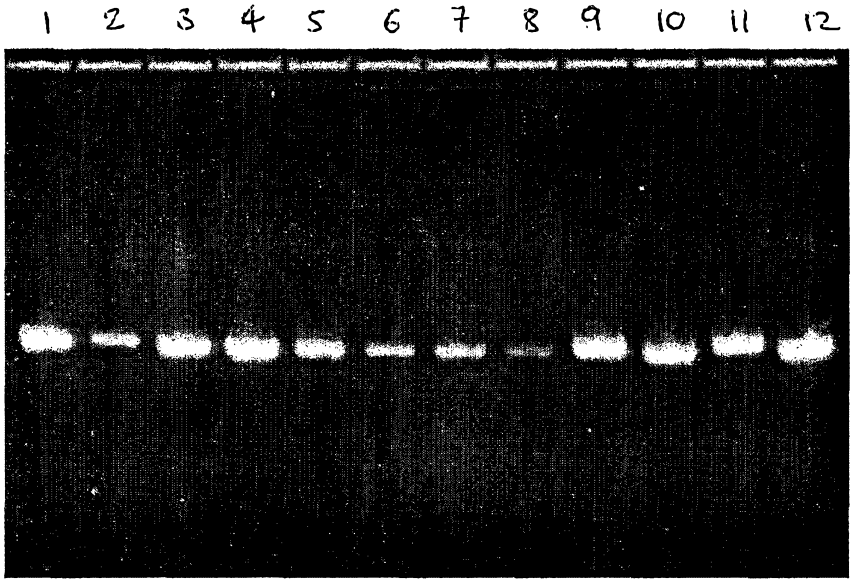


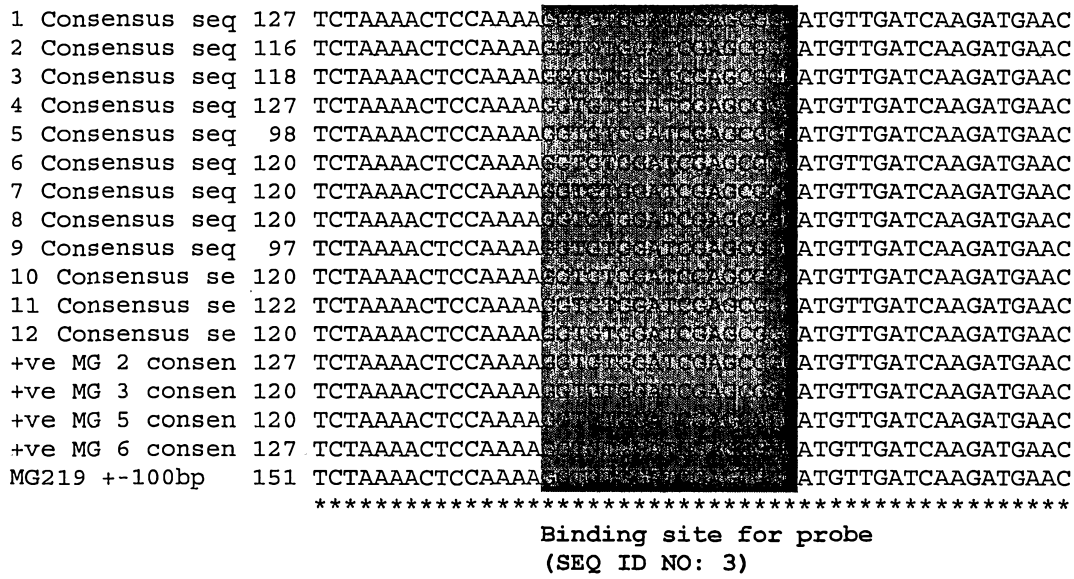
Figure 6



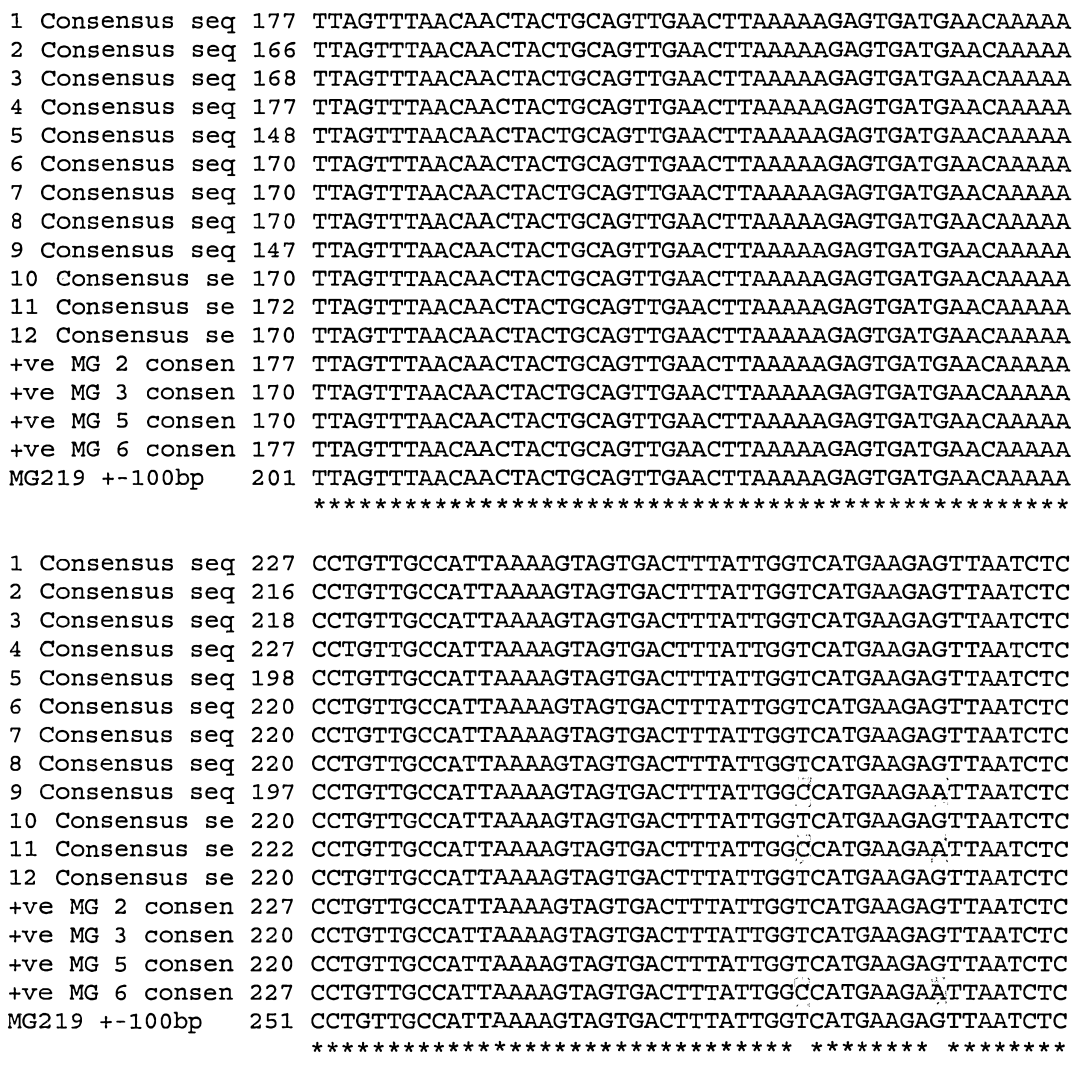
```
1 Consensus seq                                  CCACTTAACTTTATTACCCGTCCTGT  26
2 Consensus seq                                        TATTACCCGTCCTGT  15
3 Consensus seq                                     TTTATTACCCGTCCTGT  17
4 Consensus seq                                  CCACTTAACTTTATTACCCGTCCTGT  26
5 Consensus seq                                                          0
6 Consensus seq                                        ACTTTATTACCCGTCCTGT  19
7 Consensus seq                                        ACTTTATTACCCGTCCTGT  19
8 Consensus seq                                        ACTTTATTACCCGTCCTGT  19
9 Consensus seq                                                          0
10 Consensus seq                                       ACTTTATTACCCGTCCTGT  19
11 Consensus seq                                    CTACTTTATTACCCGTCCTGT  21
12 Consensus seq                                       ACTTTATTACCCGTCCTGT  19
+ve MG 2 consen                                  CCACTTCACTTTATTACCCGTCCTGT  26
+ve MG 3 cons                                          ACTTTATTACCCGTCCTGT  19
+ve MG 5 consen                                        ACTTTATTACCCGTCCTGT  19
+ve MG 6 consen                                  CCACTTAACTTTATTACCCGTCCTGT  26
MG219 +-100bp       TTGAAAAAACAACAGAAGAAAAAACCACTTAACTTTATTACCCGTCCTGT  50
                                  Binding site for sequencing primer
                                  (SEQ ID NO: 9)

1 Consensus seq  27 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  76
2 Consensus seq  16 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  65
3 Consensus seq  18 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  67
4 Consensus seq  27 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  76
5 Consensus seq   1    CAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  47
6 Consensus seq  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
7 Consensus seq  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
8 Consensus seq  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
9 Consensus seq   1     AAAAGTAATTTGCCACCAAAACTAAGTAAGGATGACATAGTTCATT  46
10 Consensus se  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
11 Consensus se  22 TTTCAAAAGTAATTTGCCACCAAAACTAAGTAAGGATGACATAGTTCATT  71
12 Consensus se  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
+ve MG 2 consen  27 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  76
+ve MG 3 consen  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
+ve MG 5 consen  20 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT  69
+ve MG 6 consen  27 TTTCAAAAGTAATTTGCCACCAAAACTAAGTAAGGATGACATAGTTCATT  76
MG219 +-100bp    51 TTTCAAAAGTAATTTGCCACCGAAACTAAGTAAGGATGACATAGTTCATT 100
                    ***************** ****************************
                                                Binding site for forward primer
                                                (SEQ ID NO: 4, 8 or 31)

1 Consensus seq  77 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 126
2 Consensus seq  66 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 115
3 Consensus seq  68 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 117
4 Consensus seq  77 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 126
5 Consensus seq  48 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA  97
6 Consensus seq  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
7 Consensus seq  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
8 Consensus seq  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
9 Consensus seq  47 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA  96
10 Consensus se  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
11 Consensus se  72 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 121
12 Consensus se  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
+ve MG 2 consen  77 CACACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 126
+ve MG 3 consen  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
+ve MG 5 consen  70 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 119
+ve MG 6 consen  77 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 126
MG219 +-100bp   101 CGCACCAGTTACTTGAAAAAAATACCCATAATGAATAGTGATAGTGA 150
                    **** ******************************************
                    Start
```



Binding site for probe
(SEQ ID NO: 3)



```
1 Consensus seq 277 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 326
2 Consensus seq 266 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 315
3 Consensus seq 268 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 317
4 Consensus seq 277 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 326
5 Consensus seq 248 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 297
6 Consensus seq 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
7 Consensus seq 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
8 Consensus seq 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
9 Consensus seq 247 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGACCAAC 296
10 Consensus se 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
11 Consensus se 272 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGACCAAC 321
12 Consensus se 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
+ve MG 2 consen 277 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 326
+ve MG 3 consen 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
+ve MG 5 consen 270 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 319
+ve MG 6 consen 277 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGACCAAC 326
MG219 +-100bp   301 TGTTCCAGTTTTACTAATCCCAACCCCTGTTGTTAAAGAGATTGATCAAC 350
                    ********************************************* ****
```

Binding site for reverse primer
(SEQ ID NO: 5)

```
1 Consensus seq 327 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 376
2 Consensus seq 316 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 365
3 Consensus seq 318 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 367
4 Consensus seq 327 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 376
5 Consensus seq 298 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 347
6 Consensus seq 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
7 Consensus seq 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
8 Consensus seq 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
9 Consensus seq 297 CAGTAGTTATTACTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 346
10 Consensus se 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
11 Consensus se 322 CAGTAGTTATTACTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 371
12 Consensus se 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
+ve MG 2 consen 327 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 376
+ve MG 3 consen 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
+ve MG 5 consen 320 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 369
+ve MG 6 consen 327 CAGTAGTTATTACTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 376
MG219 +-100bp   351 CAGCAGTTATTCCTCCAGTTAAAGCAAAACCAAAAGCAACTAAAAAGAAA 400
                    *** ******* ************************************

1 Consensus seq 377 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 426
2 Consensus seq 366 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 415
3 Consensus seq 368 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 417
4 Consensus seq 377 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 426
5 Consensus seq 348 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 397
6 Consensus seq 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
7 Consensus seq 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
8 Consensus seq 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
9 Consensus seq 347 ACTCCTGTTAAATCAAAACCAACTAATAAATCAACTAAACAAACAAAACC 396
10 Consensus se 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
11 Consensus se 372 ACTCCTGTTAAATCAAAACCAACTAATAAATCAACTAAACAAACAAAACC 421
12 Consensus se 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
+ve MG 2 consen 377 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 426
+ve MG 3 consen 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
+ve MG 5 consen 370 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 419
+ve MG 6 consen 377 ACTCCTGTTAAATCAAAACCAACTAATAAATCAACTAAACAAACAAAACC 426
MG219 +-100bp   401 ACTCCTGTTAAATCAAAACCAACTAGTAAATCAACTAAACAAACAAAACC 450
                    ************************ ************************
```

```
1 Consensus seq 427 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 476
2 Consensus seq 416 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 465
3 Consensus seq 418 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 467
4 Consensus seq 427 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 476
5 Consensus seq 398 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 447
6 Consensus seq 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
7 Consensus seq 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
8 Consensus seq 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
9 Consensus seq 397 TAAACAACCCAAACCCAAATCAAAACAAGTTCAAAAAACCAAAGCTAAAC 446
10 Consensus se 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
11 Consensus se 422 TAAACAACCCAAACCCAAATCAAAACAAGTTCAAAAAACCAAAGCTAAAC 471
12 Consensus se 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
+ve MG 2 consen 427 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 476
+ve MG 3 consen 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
+ve MG 5 consen 420 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 469
+ve MG 6 consen 427 TAAACAACCCAAACCCAAATCAAAACAAGTTCAAAAAACCAAAGCTAAAC 476
MG219 +-100bp   451 TAAACAATCCAAGCCCAAATCAAAACAAGTTCAACAAACCAAAGCTAAAC 500
                    ******* **** ********************** ***************


1 Consensus seq 477 CAA-----------------------------CCCAAATTCAAACAAAA 496
2 Consensus seq 466 CAA-----------------------------CCCAAATTCAAACAAAA 485
3 Consensus seq 468 CAA-----------------------------CCCAAATTCAAACAAAA 487
4 Consensus seq 477 CAA-----------------------------CCCAAATTCAAACAAAA 496
5 Consensus seq 448 CAA-----------------------------CCCAAATTCAAACAAAA 467
6 Consensus seq 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
7 Consensus seq 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
8 Consensus seq 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
9 Consensus seq 447 CAAAAGCAACTAAACAAACCAAAACAAAGCCAACCCAAATTCAAACAAAA 496
10 Consensus se 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
11 Consensus se 472 CAAAAGCAACTAAACAAACCAAAACAAAGCCAACCCAAATTCAAACAAAA 521
12 Consensus se 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
+ve MG 2 consen 477 CAA-----------------------------CCCAAATTCAAACAAAA 496
+ve MG 3 consen 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
+ve MG 5 consen 470 CAA-----------------------------CCCAAATTCAAACAAAA 489
+ve MG 6 consen 477 CAAAAGCAACTAAACAAACCAAAACAAAGCCAACCCAAATTCAAACAAAA 526
MG219 +-100bp   501 CAA-----------------------------CCCAAATTCAAACAAAA 520
                    ***                              ****************


1 Consensus seq 497 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 545
2 Consensus seq 486 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 534
3 Consensus seq 488 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 536
4 Consensus seq 497 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 545
5 Consensus seq 468 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 516
6 Consensus seq 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
7 Consensus seq 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
8 Consensus seq 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
9 Consensus seq 497 AAAAAGCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 546
10 Consensus se 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
11 Consensus se 522 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 570
12 Consensus se 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
+ve MG 2 consen 497 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 545
+ve MG 3 consen 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
+ve MG 5 consen 490 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 538
+ve MG 6 consen 527 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 575
MG219 +-100bp   521 AAAA-GCAATAAAAAAACCAGATCT[TCTGGTTTTTTTAGTGTTAACA 569
                    **** ******************************************
                                                 Stop
```

```
1 Consensus seq 546 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 595
2 Consensus seq 535 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 584
3 Consensus seq 537 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 586
4 Consensus seq 546 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 595
5 Consensus seq 517 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 566
6 Consensus seq 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGG   586
7 Consensus seq 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGG   586
8 Consensus seq 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 588
9 Consensus seq 547 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGG   594
10 Consensus se 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTG    585
11 Consensus se 571 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACC      615
12 Consensus se 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 588
+ve MG 2 consen 546 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 595
+ve MG 3 consen 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 588
+ve MG 5 consen 539 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 588
+ve MG 6 consen 576 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 625
MG219 +-100bp   570 ACAGTAGTTTTAACTAAAAACTGGAAAGGAATAGGAAGAACTACCTGGGG 619
                    *****************************************
```

Binding site for sequencing Primer (SEQ ID NO: 20)

```
1 Consensus seq 596 TTAA                    599
2 Consensus seq 585 T                       585
3 Consensus seq 587 T                       587
4 Consensus seq 596 T                       596
5 Consensus seq 567 TT                      568
6 Consensus seq 587                         586
7 Consensus seq 587                         586
8 Consensus seq 589 TTAAT                   593
9 Consensus seq 595                         594
10 Consensus se 586                         585
11 Consensus se 616                         615
12 Consensus se 589                         588
+ve MG 2 consen 596 TTAAT                   600
+ve MG 3 consen 589                         588
+ve MG 5 consen 589 TTAAT                   593
+ve MG 6 consen 626 T                       626
MG219 +-100bp   620 TTAATC                  647
```

Figure 7

Figure 8

Similarity Scores (%)

	1	2	3	4	5	6	7	8	9	10	11	12	+veMG2	+veMG3	+veMG5	+veMG6	G37
1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
9	93.1	93.1	93.1	93.1	93.1	93.1	93.1	93.1	100.0	93.1	99.8	93.1	92.9	93.1	93.1	99.8	93.1
10	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
11	93.2	93.2	93.2	93.2	93.2	93.2	93.2	93.2	99.8	93.2	100.0	93.2	93.1	93.2	93.2	100.0	93.2
12	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
+ve MG 2	99.8	99.8	99.8	99.8	99.8	99.8	99.8	99.8	92.9	99.8	93.1	99.8	100.0	99.8	99.8	93.1	99.8
+ve MG 3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
+ve MG 5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0
+ve MG 6	93.2	93.2	93.2	93.2	93.2	93.2	93.2	93.2	99.8	93.2	100.0	93.2	93.1	93.2	93.2	100.0	93.2
G37	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	93.1	100.0	93.2	100.0	99.8	100.0	100.0	93.2	100.0

Identity Scores (%)

The Sequence Listing for this application has been submitted to WIPO in electronic format and can be obtained upon request from the International Bureau or found on the WIPO website at http://www.wipo.int/pctdb/en/sequences/