Optimization of a nucleic acid-based reporter system to detect *Mycobacterium tuberculosis* antibiotic sensitivity

Running Title: SGM detection of antibiotic susceptibility

Matthew C. Mulvey,1,2,# Margaret Lemmon,1,3 Stephanie Rotter,1,4 Jonathan Lees,1,5 Leo Einck,¹ and Carol A. Nacy¹

¹Sequella, Inc., 9610 Medical Center Drive, Suite 200, Rockville, MD 20850 USA

²BeneVir Biopharm, Inc. 9610 Medical Center Drive, Suite 200, Rockville, MD 20850 USA

³Present Address: The Geneva Foundation 917 Pacific Ave. Suite 600 Tacoma, WA 98402

⁴Present Address: Accenture 560 Mission Street, Suite 1200 San Francisco, CA 94105

⁵Present Address: PlantVax, Inc. 9430 Key West Ave # 120, Rockville, MD 20850

#To whom correspondence should be addressed: matthewmulvey@sequella.com

Key words:

Antibiotic Resistance
Antibiotic Susceptibility Testing
Bacteriophage
Extensively-drug-resistant
Multi-drug-resistant
Surrogate Marker Locus
Surrogate Marker Locus Generation Module
Tuberculosis
Abbreviations:

CFU: Colony Forming Units
EMB<sup>r</sup>: ethambutol resistant
MDR-TB: multi-drug-resistant tuberculosis
Mtb: *Mycobacterium tuberculosis*
PCR: Polymerase Chain Reaction
PFU: plaque forming units
PI: post-infection
RIF<sup>r</sup>: rifampin resistant
RT-PCR: Reverse Transcriptase Polymerase Chain Reaction
SGM: Surrogate Marker Locus Generation Module
SML: Surrogate Marker Locus
TB: Tuberculosis
XDR-TB: Extensively-drug-resistant tuberculosis
ABSTRACT

We previously reported development of a prototype antibiotic sensitivity assay to detect drug-resistant *Mycobacterium tuberculosis* using infection by mycobacteriophage to create a novel nucleic acid transcript, a surrogate marker of mycobacterial viability, detected by reverse transcriptase-polymerase chain reaction (RT-PCR). (1) This assay detects antibiotic resistance to all drugs, even drugs for which the resistance mechanism is unknown or complex: it is a phenotypic readout using nucleic acid detection. In this report, we describe development and characteristics of an optimized reporter system that directed expression of the RNA cyclase ribozyme, which generated circular RNA through an intra-molecular splicing reaction and led to accumulation of a new nucleic acid sequence in phage-infected bacteria. These modifications simplified the assay, increased the limit of detection from $10^4$ to $<10^2$ *Mycobacterium tuberculosis* cells, and correctly identified the susceptibility profile of *Mycobacterium tuberculosis* strains exposed for 16 hr to either first-line or second-line anti-tubercular drugs. In addition to phenotypic drug resistance or susceptibility, the assay reported streptomycin minimal inhibitory concentrations and clearly detected 10% drug-resistant cells in an otherwise drug-susceptible population.
INTRODUCTION

Multi-drug resistant Tuberculosis (MDR-TB) is caused by *Mycobacterium tuberculosis* (Mt) strains resistant to at least two of the drugs, isoniazid (INH) and rifampin (RIF), in the recommended 4-drug TB treatment regimen.(2) The World Health Organization (WHO) estimates prevalence of MDR-TB at approximately 50 million people worldwide,(3) expanding by nearly 500,000 new cases each year.(4) A more disquieting development is the increase in and global distribution of extremely drug-resistant TB (XDR-TB) (Mt resistant to INH, RIF and key second-line drugs).(5) Early recognition of M/XDR-TB through accurate and sensitive AST promises to improve patient outcomes and assist in TB control efforts.

Due to the notoriously slow doubling time of Mt, standard AST using culture takes several weeks to months. Culture-based assays are the gold standard for AST, however, because they measure every type of drug susceptibility and do not require knowledge of the genetics underlying resistance. But they are slow or too insensitive to interrogate effects of antibiotics on small numbers of Mt present in most clinical samples. Rapid molecular diagnostic tests such as Xpert MTB/RIF (Cepheid) and GenoType MTBDR (Hain Lifescience) are the emerging standard for rapid INH- or RIF-resistant Mt detection, markers of MDR-TB, because their detection thresholds from patient specimens are comparable to culture.(6, 7)

These tests identify Mt DNA in a patient sample and detect common mutations that confer resistance to INH (GenoType MTBDR) and/or RIF (GenoType MTBDR and Xpert MTB/RIF) in 2 hr (MTB/RIF) or 2 days (GenoType MTBDR) from initiation of the test. However, while these assays identify MDR-TB, they do not detect susceptibility to the other first- or second-line drugs. They are an excellent epidemiologic tool and an indicator of what drugs not to use, but
they do not function as an AST should, guiding clinicians to the drugs that will be effective for 
 treating an individual patient.

For anti-tubercular drugs other than RIF and INH, the various genes and genetic 
mutations responsible for clinical drug resistance are either unknown or too numerous to test in 
a simple assay. Clearly, a rapid technology capable of detecting phenotypic drug resistance 
from low numbers of Mtb is needed. To achieve this goal, we designed a nucleic acid detection 
system that measures the phenotypic effects of antimicrobials. This strategy combines the best 
attribute of nucleic acid detection methods, namely the ability to rapidly detect low numbers of 
cellular nucleic acids, with the best attribute of phenotypic assays, namely the ability to detect 
the biological effects antimicrobials exert on susceptible microorganisms.

The reporter system has 2 functional cassettes: 1) the SP6 RNA polymerase (SP6Pol) 
gene under transcriptional control of a promoter active in *Mycobacteria* and 2) the consensus 
SP6Pol promoter that directs expression of an SP6Pol-dependent transcript.(1) The reporter 
system is incorporated into a mycobacteriophage for delivery into Mtb and the SP6Pol 
dependent transcript is detected by RT-PCR in infected cells. In our previous publication 
describing the prototype system, the assay detected ≥10⁴ Mtb cells and correctly identified 
drug-resistant Mtb strains. The SP6Pol-dependent transcript is called the Surrogate Marker 
Locus (SML) since it is a nucleic acid that functions as a surrogate marker of bacterial 
metabolic capacity. The reporter system is called the SML Generation Module (SGM), and the 
reporter phage encoding the prototype SGM is phSGM1. The prototype SML expressed in 
cells infected with phSGM1 is a linear RNA copy of its cognate DNA locus in the phage 
genome. Because this RNA transcript is not present in the input phage, synthesis of the SML 
indicates the presence of metabolically active Mtb. Detection of this linear SML RNA, however,
is laborious because phage DNA must be degraded prior to RT-PCR. Furthermore, the prototype phSGM1 assay could not detect less than $10^4$ colony forming units (CFU) of Mtb.

In this report, we describe phSGM2, a second-generation, improved version of phSGM1. The SGM in phSGM2 was optimized for function in *Mycobacteria*, encoded the RNA cyclase ribozyme under SP6Pol transcriptional control, and generated a circular SML RNA (cSML) resulting in a new nucleic acid sequence distinct from its cognate DNA locus. This simplified the assay and improved both limits of Mtb detection and ability to detect drug resistant Mtb strains, including clinical isolates, by themselves, as well as when seeded into drug-sensitive populations.

**MATERIALS AND METHODS**

**Propagation of bacterial strains and phages.** *M. smegmatis* mc²155 and mc²4502(8) were obtained from William Jacobs, Albert Einstein College of Medicine. All *Mycobacteria* were maintained in 7H9 broth supplemented with Middlebrook OADC and 0.01% Tween-80. Both the RIF-resistant (RIF<sup>R</sup>) and Ethambutol-resistant (EMB<sup>R</sup>) Mtb strains used in this study were created and isolated from H37Rv by Sequella, Inc., and maintained as for H37Rv. The MIC of RIF<sup>R</sup> Mtb to RIF and EMB<sup>R</sup> Mtb to EMB were >32 μg/ml for each. Both resistant strains were susceptible to all other anti-TB drugs tested. De-identified clinical isolates tested in this study were previously obtained from the State of Maryland TB Lab and are maintained at Sequella. The drug susceptibility profile of each strain was determined by BACTEC 460 or agar proportion method (data not shown). Recombinant mycobacteriophages phSGM2 and phP<sub>left-Cyc</sub> were propagated using mc²4502 as described.(8)

**Recombinant mycobacteriophage construction and production.** The SGM was synthesized de novo by GenScript, USA. To create phSGM2, we integrated the SGM into a
mycobacteriophage TM4 phasmid between nucleotides 47314-50108 of the TM4 genome. SGM integration was performed by GeneBridges, GmBH using Red/ET recombineering. To create ph\(_{\text{left}}\)-Cyc, we amplified by PCR the region in the SGM downstream of the SP6 promoter and upstream of the SP6MCS-T7-rrnC termination region. This product was then fused to the P\(_{\text{left}}\) promoter and integrated into the mycobacteriophage TM4 phasmid between nucleotides 47314-50108 of the TM4 genome using Red/ET recombineering.

Phasmid DNA was then isolated and electroporated into mc\(^2\)4502 to create infectious phage, as described(9). Phage stocks were prepared by elution from confluent and lysed mc\(^2\)4502 agarose overlays using MP Buffer (50mM Tris, 150mM NaCl, 10mM MgCl\(_2\), 2mM CaCl\(_2\), pH=7.6) followed by low speed centrifugation to pellet insoluble debris and filtration to sterilize the preparation. Phage stocks were adjusted to \(10^{10}\) plaque-forming units (PFU)/ml in MP buffer, treated with 50ng RNase A per ml for 2 hr at 37°C to degrade cSML produced during phage growth on mc\(^2\)4502 overlays, and then stored at 4°C.

**Phage infection.** *Mycobacteria* were washed 3X with 7H9 broth supplemented with OADC (7H9-A) to remove Tween-80. Mycobacterial suspensions (0.2ml) were then added to the wells of a 48-well tissue culture dish and incubated at 37°C for for 2 hr (*M. smegmatis* mc\(^2\)155) or a minimum of 16 hr (Mtb). Phage (50μl) in MP Buffer was then added and infected cells were incubated at 37°C.

**RNA Purification.** Total RNA was isolated by adding 0.75ml of Trizol-LS containing 1.3μg/ml polyI:C to 250μl of sample. The mixture was then transferred to a Lysing Matrix B tube (MP Biomedicals) and cells were disrupted by 2 50-second rounds of bead beating on a Mini-Bead Beater (BioSpec Products) set at 4800 oscillations per minute. Samples were cooled on ice for at least 1 min between bead beating cycles. Chloroform:isoamyl alcohol 24:1 (0.2ml) was then added, the aqueous fraction combined with an equal volume of 70% ethanol and the
final suspension added to an RNeasy silica spin column (Qiagen). RNA was then purified according to the manufacturer’s directions and eluted in 50μl RNase-free water containing 4U RNaseOUT (Life Technologies).

One-step real time RT-PCR to detect cSML. A 20μl reaction mix was prepared that contained 7.4μl of purified total RNA, 10μl EXPRESS SuperScript® qPCR SuperMix Universal, 2μl EXPRESS SuperScript® Mix for One-Step qPCR, and 200nM of oligonucleotides, specifically, P1 (5’-GCCTCTTGACAATAGGCGA3’), P2 (5’-CAATGAGTTCTGCTGCTTTTG-3’), and a molecular beacon (5’-FAM-CGCACGATGTAAAGCCCGCTGCTGCG-DABCYL-3’). EXPRESS One-Step reagents were purchased from Life Technologies and oligonucleotides were synthesized by Eurofins/Operon MWG and maintained in 10mM Tris-HCl pH=8.0/0.1mM EDTA pH=8.0. One-Step RT-PCR was performed on a PikoReal 24-well real time PCR system (Thermo Fisher Scientific) using the following conditions: 50°C for 15 min, 94°C 3 min, then 45 cycles of 94°C (30s), 57°C (30s), and 72°C (30s). FAM data acquisition was performed during the 57°C extension step.

RT-PCR to detect cSML and 16SrRNA using SYBR Green. A 20μl reaction mix contained 5μl purified total RNA, Superscript III RT (Life technologies), and 2μM of P1 or 16S Reverse (5’-GCGACGCTCAGTTAAGCCGTG-3’) oligonucleotides. Reactions were incubated at 42°C for 30 min and heat inactivated at 70°C for 15 min. Two μl of RT reaction was then added to a 20μl Power SYBR Green PCR Master Mix (Life Technologies) containing 200nM of P2 or 16SSForward (5’-GCTTAGCGGTGGATGAGCC-3’) oligonucleotides. PCR was performed using the following conditions: 94°C 3 min, then 45 cycles of 94°C (30s), 57°C (30s), and 72°C (30s). SYBR Green data acquisition was performed during the 57°C extension step.
Determining the dynamic range of cSML detection by one-step RT-PCR. *In vitro* transcribed cSML was generated by linearizing 100ng of plasmid pSP6Pro-Cyc with HindIII followed by addition to an *in vitro* transcription reaction containing 1U/μl purified SP6 RNA polymerase (New England Biolabs), 40mM Tris-HCl pH=7.9, 10mM DTT, 2mM spermidine, 4mM each rNTP, 8mM MgCl₂, and 0.8U/μl RNaseOUT. The reaction was incubated at 40°C for 1 hr. DNA was then digested by adding RNase-free DNaseI and DNaseI reaction buffer (New England Biolabs) and incubating at 37°C for 2 hr. EDTA was then added to a final concentration of 18mM and the sample heat-inactivated at 75°C for 10 min. Finally, duplicate 10-fold serial dilutions of *in vitro*-transcribed cSML were made in RNase-free water and added to a one-step RT-PCR reaction. The average cycle detection threshold for each dilution was plotted to determine the dynamic range of the one-step RT-PCR cSML detection assay, which is described by the equation \( y = (10^9)^{-0.557X} \) with \( X \) = cycle detection threshold and \( y \) = the relative level of cSML in a sample.

**Determination of limit of detection.** H37Rv or mc²155 cells were washed and adjusted to OD600nm=0.2. Serial 10-fold dilutions of cells were prepared in 7H9-A media. A 0.2ml quantity of each dilution was added to duplicate wells of a 48-well dish and also plated on duplicate 7H10 agar dishes to determine the number of CFU. For detection of mc²155, cells in 48-well culture dishes were incubated for 2 hr at 37°C and infected with 50μl of a 10⁷ PFU/ml stock of phSGM2 for 4 hr. For detection of H37Rv, cells in 48-well culture dishes were incubated for 16 hr at 37°C and infected with 50μl of a 10⁶ PFU/ml stock of phSGM2 for 4 hours. RNA was then purified, cSML detected by real time RT-PCR, and relative levels of cSML in samples calculated.

**Antibiotic Susceptibility Testing.** Washed mycobacterial cells were adjusted to OD600nm=0.02 in 7H9-A media containing antibiotics at the following concentrations: 1μg/ml...
RIF, 0.1 μg/ml INH, 1 μg/ml streptomycin (STR), 5 μg/ml EMB, 500 μg/ml p-nitrobenzoic acid (PNB), 0.25 μg/ml moxifloxacin (MOX), 5 μg/ml ethionamide (ETH), 2 μg/ml para-amino salicylic acid (PAS), 4 μg/ml kanamycin (KAN), 1 μg/ml amikacin (AMK), 1.6 μg/ml SQ109, and 40 μg/ml cycloserine (CS). Cells (0.2 ml) were then added to wells of a 48-well dish and incubated at 37 °C for either 16 or 40 hr. The phSGM2 diluted to 10^7 PFU/ml in MP Buffer (50 μl) was then added and incubated at 37 °C for 4 hr. Total RNA was then purified using Trizol-LS and bead beating and cSML and/or 16S rRNA amplified, detected, and relative levels of cSML in samples calculated.

To determine RIF and STR MIC using Trek Sensititre MycoTB plates (Thermo Fisher Scientific), we added 100 μl of washed RIF^R cells at OD600nm=0.02 to the wells and incubated the plate at 37 °C for 16 hr and then infected the wells with 25 μl of phSGM2 diluted to 10^7 PFU/ml in MP Buffer for 8 hr. We added 100 μl of 100% ethanol to the wells and the mixture was transferred to a microfuge tube containing 200 μl 100% ethanol and incubated at RT for 15 min to inactivate Mtb. Buffer RLT (425 μl, Qiagen) was then added and the mixture applied to an RNeasy silica spin column. Total RNA was then purified and cSML amplified, detected, and relative levels of cSML in samples calculated.

Susceptibility to an antibiotic was determined if drug treatment resulted in >10-fold reduction in cSML levels compared to the untreated control. Conversely, resistance was scored if drug treatment resulted in =10-fold reduction in cSML levels compared to the untreated control. These cutoff values were based on the extensive work performed by the Jacobs group to evaluate rapid AST using recombinant luciferase reporter phages.(10)

RESULTS

Detection of Mycobacteria by phSGM2. Figure 1A and B describes the overall construction of the various components of phSGM2, and the actions and order of actions...
following infection of Mtb with the recombinant phage that results in the detection of viable bacterial cells. To determine the lowest number of cells detected by phSGM2, we made serial dilutions of *M. smegmatis* strain mc²155 or *M. tuberculosis* strain H37Rv (Figure 1C). The dilutions were plated to determine the number of CFU and also infected with phSGM2. Uninfected 10⁵ CFU and phage alone served as controls. At 4 hr post-infection (PI), total RNA was purified and cSML measured by RT-PCR. The cSML was not detected in samples containing approximately 10⁵ cfu of either mc²155 or H37Rv (Figure 1C), but was observed in phage alone controls. Despite the presence of cSML in the phage alone controls, cSML generation from as few as 18 CFU mc²155 or 94 CFU H37Rv infected with phSGM2 were detected with statistical significance (Figure 1C).

**AST of Mtb using phSGM2 reporter phage.** To evaluate the ability of phSGM2 to determine Mtb antibiotic susceptibility, we left drug susceptible H37Rv cells untreated or treated them with front line anti-TB drugs RIF, INH, STR, or EMB. Cells were also treated with PNB, which exhibits anti-mycobacterial activity only against members of the TB complex and is used to verify that a sample does not contain non-tuberculous *Mycobacteria*. After either 16 hr or 40 hr of antibiotic exposure, the cells were infected with phSGM2 for 4 hr. Total RNA from each sample was then purified and cSML detected by RT-PCR. The level of cSML in untreated cells infected with phSGM2 was approximately 1000-fold higher than the phage alone controls. The level of cSML in infected cells exposed to antibiotics for either 16 hr or 40 hr was reduced to background levels (Figure 2A). Using >10-fold reduction in drug treated samples compared to untreated controls as the cutoff for drug susceptibility, these data demonstrate that phSGM2 could report susceptibility to all first line anti-TB antibiotics, excluding PZA (not shown), after less than 1 day of drug exposure. Furthermore, the dynamic range of the assay was approximately 1000-fold between drug treated and untreated samples. The phSGM2 also reported susceptibility to all second line anti-TB drugs tested (Figure 2B). To determine if the...
differences in cSML levels between untreated and drug-treated samples was due to effects on 
cSML generation or to global effects on cellular RNA, or differences in the number of cells 
added to each sample, we performed RT-PCR to detect cSML and 16S rRNA using purified 
total RNA from the samples in Figure 2A (Table 1). In samples treated with antibiotics, cSML 
levels were reduced 12-15 cycles compared to the untreated control; whereas, 16SrRNA was 
reduced by, at most, 4.4 cycles (RIF treated sample). This demonstrated that the 
approximately 1000-fold differences in cSML levels observed between untreated and drug 
treated cells infected with phSGM2 were not due to loading errors or pleotropic effects of 
antibiotics on cellular RNA.

To determine if phSGM2 reported drug resistance, we exposed two H37Rv variants to 
nothing (No Drug) or to RIF, EMB, or PNB. One variant was susceptible to EMB and PNB, but 
resistant to RIF (RIFR) and the other was susceptible to RIF and PNB, but resistant to EMB 
(EMBR). After 16 hr of drug exposure, Mtb were infected with phSGM2 for 4 hr, total RNA was 
purified, and cSML detected by RT-PCR (Figure 2D). cSML generation in untreated RIFR and 
EMBR cells infected with phSGM2 was robust while exposure to PNB reduced cSML 
generation between 100-1000 fold, demonstrating that the organisms were Mtb. Treatment of 
the EMB susceptible RIFR variant with EMB reduced cSML generation similarly to PNB. 
However, treatment with RIF did not significantly reduce cSML generation because the variant 
was resistant to RIF. Finally, treatment of the RIF susceptible variant EMBR with RIF reduced 
cSML generation > 100-fold. However, treatment with EMB did not significantly reduce cSML 
generation because the variant was resistant to EMB. These data demonstrate that cSML 
generation in phSGM2 infected cells could identify both drug susceptibility and drug resistance.

H37Rv is a standard, widely used Mtb laboratory strain that is highly metabolic 
compared to clinical isolates, especially M/XDR-TB. To determine if phSGM2 is able to report 
drug susceptibility and resistance in clinical isolates, several INH and STR resistant clinical
isolates maintained at Sequella, Inc. were tested for susceptibility to INH and RIF (Figure 2D) or STR and KAN (Figure 2E) after 40hrs of drug exposure prior to phage infection. In all clinical isolates tested except 8330, cSML levels in the untreated control were elevated at least 100-fold over PNB background. Strain 8330 was the slowest growing of all isolates tested (MM observations) and cSML levels in the untreated control were 16-fold higher than the PNB background control (Figure 2D) reflecting the low metabolic activity of this strain. Despite the low metabolic activity of strain 8330, phSGM2 correctly reported 8330 RIF susceptibility and INH resistance. In addition, phSGM2 2 correctly reported the susceptibility profile of the other clinical strains to INH, RIF, STR and KAN.

Expression of cSML by SP6Pol is required to accurately report drug susceptibility. Synthesis of cSML in phSGM2 infected cells could be mediated by host RNA polymerase rather than SP6Pol. Furthermore, reporting of drug susceptibility by phSGM2 could simply reflect lower levels of phage gene expression by host RNA polymerase. In this situation, it would be far simpler to use wildtype phages as reporters and measure accumulation of an endogenous phage transcript. To determine if cSML synthesis by SGM encoded SP6Pol was required for accurate reporting of drug susceptibility, we constructed phPleft-Cyc (Figure 3A). Compared to phSGM2, phPleft-Cyc does not encode SP6Pol and the Pleft promoter directs host RNA polymerase to express cSML. As can be seen in Figure 3B, H37Rv cells infected with phSGM2 synthesized dramatically higher levels of cSML compared to cells infected with phPleft-Cyc. This suggests that there was an SP6Pol-mediated positive feedback loop. Specifically, a subset of transcripts initiated at the SP6 promoter by SP6Pol continued through the triplet termination unit downstream of RC2 and transcribed the SP6Pol ORF, resulting in enhanced synthesis of SP6Pol compared to that expressed by host RNA polymerase, which in turn mediated increased cSML generation. Levels of cSML in RIF-treated cells infected with either phSGM2 or phPleft-Cyc were reduced to background levels.
Furthermore, in phSGM2 infected cells treated with STR, cSML levels were reduced to background, indicating STR susceptibility. However, cSML levels were unaffected in cells infected with pHmet-Cyc and treated with STR. These findings suggest that accurate reporting of susceptibility to bacteriostatic agents that target cellular protein synthesis must couple signal generation to translation.

**phSGM2 reports STR MIC.** RifR cells were added to wells of a Sensititre MycobTB dish containing STR, RIF, or excipient control to determine the MIC of STR and RIF. After overnight incubation, cells were infected with phSGM2 and cSML generation measured by RT-PCR. All STR concentrations from 0.5 μg/ml and above resulted in approximately 1000-fold reduction in cSML synthesis. However, once the STR concentration fell below 0.5 μg/ml, cSML generation was indistinguishable from the untreated controls. This identifies 0.5 μg/ml as the STR MIC against this Mtb strain, which agrees with BACTEC 460 (data not shown). No MIC for RIF was identified because the strain was resistant to RIF.

**phSGM2 detected RifR Mtb in a drug susceptible population.** To determine the level of drug resistant bacteria that phSGM2 could detect, we mixed RifR cells at varying concentrations with Mtb strain H37Rv, which is susceptible to RIF. The mixtures were treated with RIF, STR, or PNB for 40 hr, infected with phSGM2, and cSML generation detected by RT-PCR (Figure 5). No significant difference was observed in cSML levels in all samples containing H37Rv or RifR Mtb exposed to STR or PNB. However, cSML generation in a population containing 10% RifR cells was elevated >10-fold above the level of cSML generated in the STR and PNB treated samples. These data demonstrate that phSGM2 is able to detect the presence of 10% drug resistant cells in an otherwise drug susceptible population.

**DISCUSSION**
We designed the SGM reporter as a bridge that connects the benefits of rapid molecular diagnostic technologies like GeneXpert with those of phenotypic culture-based tests capable of identifying all forms of drug resistance and the presence of drug-resistant bacteria in an otherwise drug-susceptible population. To this end, we demonstrated that phSGM2 encoding the second generation SGM detected <100 CFU (Figure 1C) of Mtb, which compares favorably with the reported 131 CFU limit of detection for Xpert MTB/RIF. (7) With the exception of PZA, phSGM2 accurately reported Mtb susceptibility to all antibiotics used to treat TB patients (Figures 2A & B) in less than one day. Furthermore, the assay correctly reported each drug resistant strain tested, including slow growing clinical isolates (Figures 2C-E), and clearly reported the STR MIC for an H37Rv-derived strain (Figure 4). Finally, phSGM2 was able to clearly detect the presence of 10% drug resistant cells in an otherwise drug susceptible population (Figure 5).

Other groups have reported detection of phenotypic AST using phage in combination with nucleic acid based detection: a phage amplification assay uses PCR to detect the increase in mycobacteriophage D29 genomes after low multiplicity infection of cultured Mtb. (11) Unlike the SGM assay, which produces a new nucleic acid signal during infection of Mtb, the phage amplification assay is unable to distinguish input phage genomes from those produced by phage replication. Therefore, multiple rounds of phage infection and replication are required to accumulate sufficient progeny phage genomes over the number of input phage.

Technologies that amplify and detect nucleic acids are the emerging standard for rapid pathogen identification directly from TB patient samples. Cepheid’s GeneXpert MTB/RIF test system can diagnose TB infection within two hr from initiation of the test, and with sensitivity similar to culture, which takes several days to weeks. Built into this system is simultaneous detection of RIF resistance, which is accomplished by 4 molecular beacons that span the RIF
resistance determining region (RRDR), an 81bp region in the Mtb rpoB gene responsible for nearly 100% of clinically significant resistance to RIF. The presence of a single mutation in this region disrupts binding of 1 of the 4 molecular beacons to the RRDR amplicon, which indicates RIF resistance. This is a very powerful testing platform, but it is not scalable beyond detection of RIF resistance because the number of mutations responsible for resistance to any of the other front-line and second-line drugs used to treat TB patients are far more numerous and spatially distributed throughout the Mtb genome. Furthermore, the strategy of detecting mutations by measuring the absence of molecular beacon binding requires that at least 65% of the Mtb genomes in a patient sample contain the mutation.(12) Therefore, this approach does not meet the definition of drug resistance set by the gold standard agar proportion method, which is the presence of 1% drug-resistant cells in an otherwise drug susceptible population.(13, 14)

Detection of antigen 85B mRNA (short half-life) is another method that uses nucleic acid-based detection of drug phenotypic effects. In this study, Mtb cultures treated with INH or RIF had rapid decreases in Ag85B mRNA and accurately reported drug susceptibility and resistance.(15) Also, multiplexed quantitative detection of several Mtb mRNA targets using nCounter Analysis (Nanostring Technologies) accurately reports phenotypic AST of cultured clinical strains.(16) To our knowledge, the effectiveness of either of these approaches has not yet been validated using clinical samples perhaps due to the complex mixture of viable, dying, and dead cells present in sputum samples that may complicate the relationship between drug effects and cellular mRNA turnover. Since the cSML is an entirely synthetic sequence not present in nature, interference of cellular RNA is not expected to be an issue.

Susceptibility to PZA at pH5.5, which is optimal for activity against Mtb, was detected by phSGM2 within 4 days. PZA exerts anti-mycobacterial activity by inhibiting tran-
translation,(17) which releases stalled translation elongation complexes in order to preserve pools of translationally-competent ribosomes. It is thought that trans-translation becomes an essential process under acidic conditions because a large fraction of ribosomes are stalled on mRNA templates that are acid hydrolyzed and cleaved 5' of a stop codon. Therefore, inhibition of trans-translation by PZA under these conditions could affect mycobacterial viability. This conditionally lethal nature of trans-translation inhibition may be responsible for delayed PZA effectiveness compared to other anti-TB drugs, and the longer time to detect resistance in our AST.

Perhaps most importantly for TB patients and clinicians, we demonstrated that phSGM2 detected as few as 10% drug-resistant cells in an otherwise drug-susceptible population. This opens the door to developing the AST as a method to detect real-time development of drug resistance in patients undergoing TB therapy. A rapid test with this capability would indicate pending failure of the current treatment regimen much earlier than is currently possible, and allow the patient to switch to second-line treatment while the drug resistant Mtb biomass is substantially lower than it is in patients diagnosed with drug resistant infection using current molecular or phenotypic tests. Such a test would identify patients likely to fail treatment, maximize the effectiveness of second line drugs, and lower the incidence of MDR-TB. The current assay clearly detects the presence of 0.1%-1% drug-resistant cells. There remains in the assay, however, background cSML resistant to RNase A and all ribonucleases tested, which co-purifies with infectious phSGM2 by CIM monolith chromatography and CsCl isopycnic gradient ultra-centrifugation, and becomes a substrate for RNase A after phSGM2 is incubated at 70°C for 15min (data not shown). We believe that a fraction of cSML produced during growth of phage stocks becomes incorporated inside phage capsids and is thus shielded from ribonucleases. Future efforts will focus on developing a SGM that precludes
accumulation of the cSML during growth of phage stocks in *M. smegmatis*. Removal of background could facilitate detection of ≤0.01% drug resistant bacteria.

In this report we demonstrate that SGM is a new genetic approach to detect microbial AST: it combines the sensitivity and accuracy of phenotypic AST (measures activity of any drug) with the power of nucleic acid tests to detect very few microorganisms in a matter of just hours. We are collaborating with TB clinical trial sites to assess the utility of this AST used directly on sputum samples and to determine its usefulness to identify drug resistance as it develops in patients.

**Acknowledgements:** This work was supported by NIAID grant 1R21AI096248-01A1 awarded to MM.

**REFERENCES**


Figure Legends

Figure 1. The phSGM2 design, cSML detection assay dynamic range, and detection of M. smegmatis and Mtb. (A) The mycobacteriophage TM4 genome is depicted as a solid black line at the top of the Figure. The direction of phage gene expression by host RNA polymerase is indicated by the dashed arrow above the phage genome. The site in the phage genome where the SGM is inserted is indicated and expanded. The SGM is comprised of two cassettes:

1. The first is the SP6Pol open reading frame (ORF) under transcriptional control of the mycobacteriophage L5 Pleft promoter (open star), which directs expression of SP6Pol during infection of Mycobacteria. The SP6Pol ORF is codon optimized for efficient translation in Mtb and translation initiation is directed by the Shine-Dalgarno sequence upstream of the TM4 major capsid subunit;

2. The second cassette encodes the SP6 promoter (filled star) fused to a downstream sequence encoding several functional modules. The SP6 promoter cassette is flanked
by transcription terminators. Upstream of the SP6 promoter is the *E. coli* rrnBT2
terminator (grey dot),(17) which precludes transcription of the SP6 promoter cassette by
host RNA Polymerase. We positioned three terminators downstream of the SP6
promoter cassette: phage SP6 major capsid subunit (SP6MCS - blue dot),(8) T7
(orange dot),(15) and the *E. coli* rrnC (green dot)(17) terminators. These terminators
were designed to lower the frequency at which elongating SP6Pol transcribes the
SP6Pol ORF. We also positioned several functional modules downstream of the SP6
promoter: T1 and T2 encoded sites to which primers P1 and P2 bind. These sites were
oriented such that upon binding to T1 and T2, primers P1 and P2 are unable to
generate a PCR amplification product. Adjacent to T1 and T2 were Exons A and B. RC1
and RC2 were positioned at the 5’ of Exon A and 3’ of Exon B, respectively.

RC1 and RC2 encode two halves of the RNA Cyclase (RC) ribozyme.(10) After transcription of
this locus by SP6 Pol, a single stranded RNA was synthesized that had one half of RC (RC1)
fused to the 5’ end of Exon A followed by T1, T2 and the other half of RC (RC2) fused to the 3’
end of Exon B. Once RC2 was synthesized, RC1 and RC2 interacted and formed the active
RC ribozyme, which mediated circularization of the single stranded RNA between RC1 and
RC2 by fusing Exons A and B via a splicing reaction. Fusion of Exons A and B created a new
nucleic acid sequence distinct from the cognate DNA locus in the phage genome and
constituted generation of the cSML. In addition, RC1 and RC2 fused to each other as a
byproduct. Finally, detection of the cSML was performed using P1 and P2 because splicing of
Exons A and B generated an intervening sequence between the 3’ ends of P1 and P2,
resulting in the creation of a template for RT-PCR amplification. The cSML was amplified using
a one-step, combined RT-PCR reaction and detected by a molecular beacon. (B) The dynamic
range was determined by cSML synthesis from HindIII digested pSP6Pro-Cyc by *in vitro*
transcription. Duplicate ten-fold serial dilutions were prepared and cSML was amplified: The
average cycle detection thresholds for each dilution were plotted and the best fit line through the data points is described by the equation: \( y = (10^9)^{0.557x} \) where \( x \) = cycle detection threshold and \( y \) = the relative level of cSML in a sample. This equation is used to determine the relative level of cSML present in a sample. (C) Determination of the lowest numbers of *M. smegmatis* mc2155 and Mtb H37Rv detected by the assay.

**Figure 2.** *phSGM2 accurately detected susceptibility and resistance to all classes of anti-TB drugs.* cSML generation in H37Rv cells infected by *phSGM2* and exposed to (A) frontline and (B) second line anti-TB drugs. (C) *phSGM2* detected resistance to RIF and EMB in RIF\(^R\) and EMB\(^R\) H37Rv variants. (D & E) *phSGM2* accurately reported the (D) INH and RIF and (E) STR and KAN susceptibility profiles of clinical isolates 386, 622, 7739, 8330, 8668, 9016.

**Figure 3.** *SP6Pol encoded in phSGM2 lead to increased cSML generation in infected cells and was required to report susceptibility to STR.* (A) cSML expression was controlled by host RNA polymerase in cells infected with mycobacteriophage *phP\(_{left}\)-Cyc*. (B) cSML levels in H37Rv cells infected with *phSGM2* accumulated to higher levels than in cells infected with *phP\(_{left}\)-Cyc*. Susceptibility to RIF was reported by both phages, but STR susceptibility was not reported by *phP\(_{left}\)-Cyc*.

**Figure 4.** *phSGM2 accurately reported the STR MIC of a RIF\(^R\) H37Rv variant.* RIF\(^R\) cells were added to the positive control, STR and RIF containing wells of a Trek Sensititre dish and
then infected with phSGM2. cSML generation was reduced to background levels at STR concentrations at or above the STR MIC of the RIF$^R$ strain.

Figure 5. phSGM2 detected the presence of 10% RIF$^R$ cells in an otherwise drug susceptible population. H37Rv, RIF$^R$ or mixtures of both containing 0.1%, 1% or 10% RIF$^R$ cells were either left untreated or exposed to RIF, STR, or PNB and then infected with phSGM2. A >10-fold increase in cSML generation above STR or PNB samples was detected in mixtures containing at least 10% RIF$^R$ cells and exposed to RIF.
TABLE 1. Cycle detection thresholds (C_q) for SYBR Green real time RT-PCR detection of cSML and 16S rRNA from H37Rv cells treated with front-line anti-TB drugs and infected with phSGM2. ND: Not Detected. Data from one representative replicate is presented.

<table>
<thead>
<tr>
<th>Drug</th>
<th>cSML</th>
<th>16S rRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Drug (No RT)</td>
<td>ND</td>
<td>41.95</td>
</tr>
<tr>
<td>No Drug</td>
<td>22.07</td>
<td>30.16</td>
</tr>
<tr>
<td>RIF</td>
<td>37.36</td>
<td>34.59</td>
</tr>
<tr>
<td>INH</td>
<td>34.95</td>
<td>32.20</td>
</tr>
<tr>
<td>STR</td>
<td>35.38</td>
<td>34.04</td>
</tr>
<tr>
<td>EMB</td>
<td>35.07</td>
<td>32.49</td>
</tr>
<tr>
<td>PNB</td>
<td>36.66</td>
<td>33.77</td>
</tr>
<tr>
<td>Phage Alone</td>
<td>36.41</td>
<td>ND</td>
</tr>
</tbody>
</table>
Figure 4

Graph showing the relative cSMG levels with varying concentrations of STR and RIF. The x-axis represents the concentration of STR and RIF in μg/ml, while the y-axis represents the relative cSMG levels in log scale.
Figure 5

Graph showing cSML level (relative) for different conditions:
- No Drug
- STR
- PNB
- RIF

Legend:
- H37Rv
- 0.1% RIF\(^R\)
- 1% RIF\(^R\)
- 10% RIF\(^R\)
- RIF\(^R\)