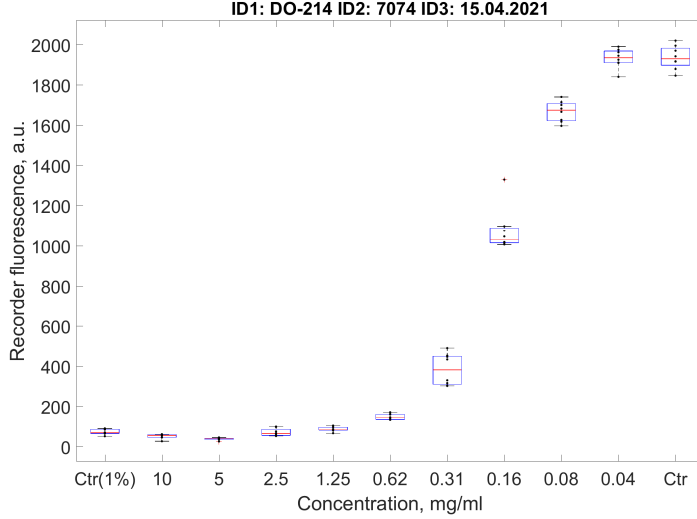


Тест экспериментальных точек

Шаг 1: первичная обработка данных, исключение первичных выбросов

Для каждого столбца двумерного массива данных \mathbf{d}_{ij} по флуоресценции строится диаграмма размаха (“ящик с усами”¹) как показано на рис. 1.



¹ box-and-whiskers plot

Рис. 1: Экспериментальные данные для различных концентраций (черные точки) и соответствующие диаграммы размаха (см. текст с описанием параметров); красные плюсы маркируют выбросы с исходных данных.

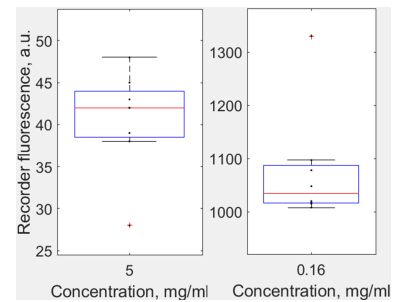
Параметры ящика стандартные для графических непараметрических методов визуализации статистических данных²: верхняя и нижняя планки ящика маркируют 1-й и 3-й квартили данных (т.е. 0.25 и 0.75 квантили при нормировке полной вероятности на единицу) q_1 и q_3 , соответственно; красная линия обозначает медиану M , границы “усов” w_{up} и w_{down} определяются стандартными соотношениями

$$w_{up} = q_3 + 1.5(q_3 - q_1), \quad (1)$$

$$w_{down} = q_1 - 1.5(q_3 - q_1) \quad (2)$$

между квартилями и интерквартильными интервалами. Точки, выходящие за пределы усов (производится проверка $\mathbf{d}_{ij} > w_{up}$ или $\mathbf{d}_{ij} < w_{down}$ для каждого столбца концентраций – указанного индексом j) считаются выбросами (*outliers*), помечаются на графике красными крестиками и удаляются из скорректированного массива данных $\tilde{\mathbf{d}}_{\{i\}j}$ для дальнейшей обработки³. При этом на графике рис. 1 согласно стандартной конвенции, усы обрезаются на соответствующем минимальном и максимальном значении из оставшихся, не являющихся выбросами. Таким образом, диаграмма размаха позволяет судить об интервале экспериментальных значений $\tilde{\mathbf{d}}_{\{i\}j}$, используемых при последующем анализе.

² J. M. Chambers, W. S. Cleveland, B. Kleiner, P. A. Tukey, Graphical Methods for Data Analysis, CRC Press (Boca Raton), 2018



³ Здесь фигурные скобки индекса $\{i\}$ означают возможное различие значения максимального индекса в различных колонках j .

Шаг 2: определение дискретной минимальной ингибирующей концентрации (МИК) для очищенной выборки

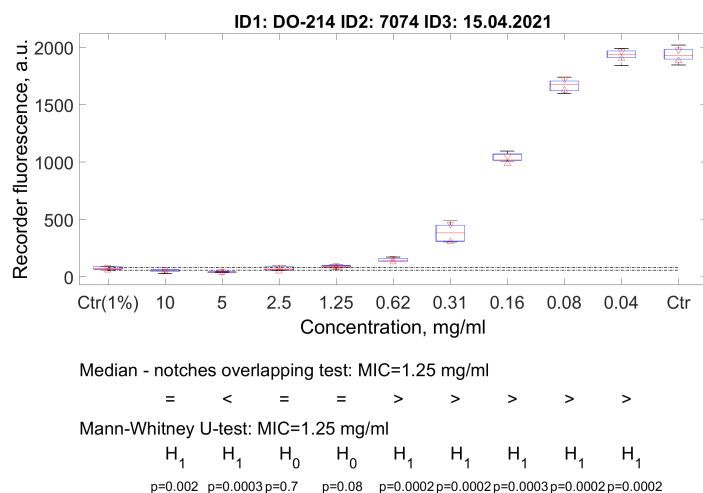
Для очищенной выборки снова проводится процедура оценки квартилей и медианы, а также положения усов по (1)–(2), но уже с новыми значениями квартилей⁴. Полученная диаграмма “ящиков с усами” показана на рис. 2.

Помимо этого, вычисляется доверительный интервал (*notches*) для каждой медианы по формулам⁵

$$\Delta_{upj} = M_j + 1.57 (q_{3j} - q_{1j}), \quad (3)$$

$$\Delta_{downj} = M_j - 1.57 (q_{3j} - q_{1j}) \quad (4)$$

Положения соответствующих значений указаны на рис. 2 как центры зеленых треугольников в вершинах, направленными вниз и вверх, соответственно. Для 1% контроля эти границы продлены на весь исследуемый интервал.



⁴ Более того, после убирания больших выбросов такой малой выборке могут появиться и новые выбросы за счет сдвига границ квартилей, однако, опять-таки в силу малости выборки, дальнейшего переопределения медианы и границ не производится.

⁵ Значение множителя 1.57 соответствует 95 % вероятности статистической незначимости различия медиан в случае нормально распределенных данных; для произвольных выборок такая непосредственная интерпретация не выполняется, однако это же значение принято как стандарт значимости различий и в этом случае.

Рис. 2: Диаграмма размаха экспериментальных данных с удаленными первичными выбросами (“ящики с усами”) на которой показаны границы доверительного интервала 1% контроля (черные штриховые линии), под которой для каждой концентрации указаны значимые совпадения или различия медиан в сравнении 1% контролем и сравнения этих выборок по критерию Манна-Уитни с указанием уровня значимости. Для обоих методов указаны соответствующие значения МИК.

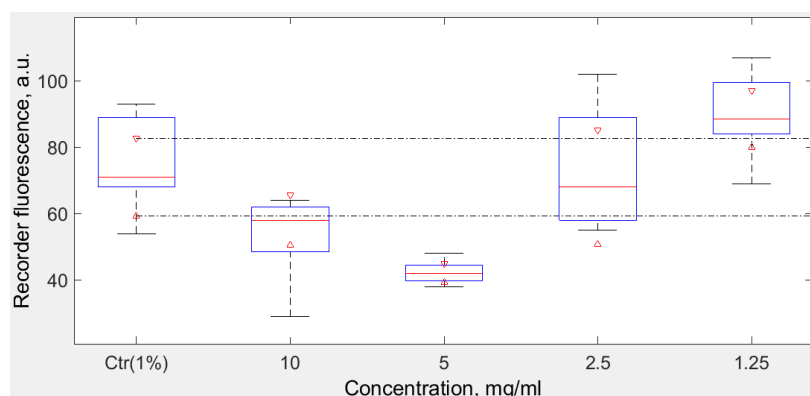


Рис. 3: Увеличенный участок больших концентраций для более наглядной демонстрации областей пересечения интервалов для 1% контроля (штриховые линии) и различных концентраций (внутри области, ограниченной треугольниками).

Для нахождения МИК проводится сравнения интервалов флуоресценции с 1% контролем с проверкой двух условий:

- (i) есть ли статистически значимое неразличие медиан

флуоресценции при данной концентрации и при 1% контроле;

- (ii) если есть статистически значимое различия медиан, то какая из них больше.

По пункту (i) автоматизированные вычисления проводится следующим образом: попарно сравниваются нижние границы доверительных интервалов для 1% контроля (Δ_{down1}) и текущей концентрации (Δ_{downj} , $j \leq 2$) и находится максимальная из них,

$$\Delta_{maxj} = \max \{ \Delta_{down1}, \Delta_{downj} \}, \quad (5)$$

верхние границы доверительных интервалов для 1% контроля (Δ_{up1}) и текущей концентрации (Δ_{upj} , $j \leq 2$) и находится минимальная из них,

$$\Delta_{minj} = \min \{ \Delta_{up1}, \Delta_{upj} \}, \quad (6)$$

после чего сравниваются результаты (5) и (6). Если $\Delta_{minj} > \Delta_{maxj}$, то различие медиан рассматривается как статистически незначимое, на рис. 2 под соответствующей концентрацией указана пометка “=”. В противном случае проверяется соотношение величин статистически различных медиан – если $M_j < M_1$, то есть медиана 1% контроля больше, то указывается пометка “<”, иначе – пометка “>”. МИК для дискретного набора концентраций определяется как концентрация, соответствующая наибольшему индексу, для которого определены “=” или “<”, то есть медиана статистически совпадает с медианой 1% контроля или меньше нее.

В качестве дополнительного теста использовано сравнение распределений экспериментальных данных для 1% контроля и различных концентраций по U-критерию Манна-Уитни (*Mann-Whitney U-test*), который служит непараметрической заменой двустороннего t-теста для средних, который определен только для нормально-распределенных данных, на случай произвольного распределенных данных и сравнивает статистические различия медиан⁶. Соответствующие пометки показаны на рис. 2: H_0 означает что сравниваемые данные соответствуют выборкам с равными медианами на уровне достоверности 5% ($p > 0.05$), а H_1 указывает на их статистически значимое различие⁷; соответствующие уровни значимости, найденные в ходе U-теста также указаны под каждым сделанными выводом. В данном случае МИК определяется как максимальный из индексов, при котором либо выполняется H_0 , либо H_1 , но с медианой при данной концентрации меньшей, чем медиана 1% контроля. На рис. 2 обе МИК совпадают.

⁶ W. J. Conover, Practical nonparametric statistics, John Wiley & Sons (New York), 1998

⁷ Различие выводов двух тестов для концентрации 10 mg/ml объясняется тем, что в методе пересечений доверительного интервала медиан отклонения от последних симметричны, но реальная выборка смещена вниз, как видно на рис. 3: хотя верхний треугольник и попадает между штриховыми линиями, сами “ящики” совершенно не перекрываются и вытянуты по отношению к медиане в противоположные стороны.

Аппроксимация кривой отклика флуоресценции на действующее вещество непрерывной кривой

Биофизическое обоснование

Дискретные значения, показанные на рис. 2, следуют зависимости уровня флуоресценции от логарифма концентрации, которую можно аппроксимировать сигмоидальной функцией. Удобно перейти к безразмерным нормированным величинам

$$m = \frac{M - M_1}{M_{contr} - M_1}, \quad (7)$$

где M_1 – значение флуоресценции 1% контроля (то есть минимально регистрируемая величина), а M_{contr} – значение флуоресценции чистой культуры (максимально регистрируемая величина).

Таким образом, $m(\ln(C))$ изменяется от 0 до 1, и в качестве ее аппроксимации возьмем сдвинутый гиперболический тангенс⁸

$$m = \frac{1}{2} \left[1 - \tanh \left(\frac{1}{2} \alpha (\ln(C) - \ln(C_{0.5})) \right) \right], \quad (8)$$

где α и $C_{0.5}$ – константы, подлежащие нахождению.

Формулу (8) можно преобразовать⁹ к виду

$$m(C) = \frac{C_{0.5}^\alpha}{C_{0.5}^\alpha + C^\alpha}, \quad (9)$$

который, в свою очередь, можно записать в эквивалентной форме

$$m(C) = 1 - \frac{C^\alpha}{C_{0.5}^\alpha + C^\alpha}. \quad (10)$$

Так как измерения флуоресценции проводятся через достаточно короткое время после введения вещества, то есть влиянием емкости среды можно пренебречь, что динамика роста бактериальной популяции N удовлетворяет уравнению Мальтуса

$$\frac{dN}{dt} = r(C)N, \quad (11)$$

где $r(C)$ – коэффициент роста, зависящий от концентрации введенного вещества; в случае контроля $r = r_0 = \text{const}$.

Для малого промежутка времени Δt производную можно заменить конечной разностью

$$\frac{N(t_0 + \Delta t) - N(t_0)}{r_0 N(t_0) \Delta t} \approx \frac{r(C)}{r_0} \quad (12)$$

⁸ Множитель 1/2 введен для удобства кинетической интерпретации, см. далее.

⁹ Пошаговые вычисления:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \equiv \frac{e^{2x} - 1}{e^{2x} + 1},$$

$$x = \ln(z), \quad 2x = \ln(z^2),$$

$$1 - \tanh(z) = 1 - \frac{z^2 - 1}{z^2 + 1} = \frac{2}{z^2 + 1}.$$

Так как величина

$$\frac{1}{2} \alpha (\ln(C) - \ln(C_{0.5})) = \ln \left(\frac{C^{\frac{\alpha}{2}}}{C_{0.5}^{\frac{\alpha}{2}}} \right),$$

то ее подстановка в качестве z приводит к (9).

и, в силу того, что уровень флуоресценции зависит от уровня прироста популяции, а также сравнивая (12) с (7), приходим к выводу, что $m(C) \approx r(C)/r_0$, то есть уравнение роста (11) может быть записано, с учетом (10), в виде¹⁰

$$\frac{dN}{dt} = r_0 N - r_0 \frac{C^\alpha N}{C_{0.5}^\alpha + C^\alpha}. \quad (13)$$

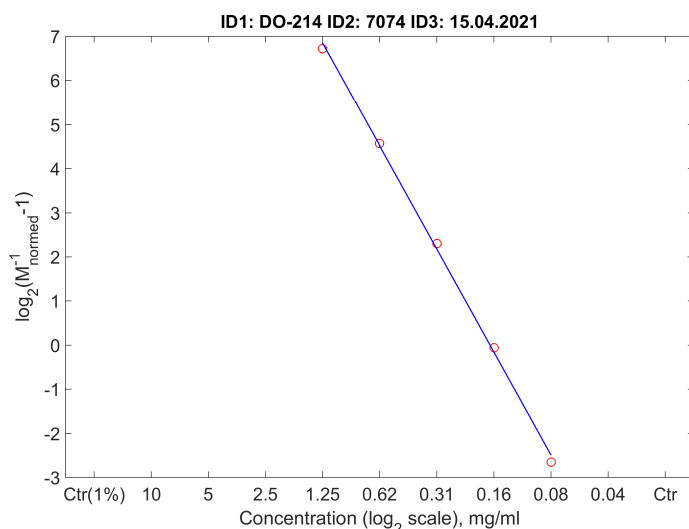
В уравнении (13) первый член соответствует невозмущенному росту, а второй – замедлению роста за счет ингибирования, имеющего точную форму фермент-субстратной ингибирующей функции Хилла с концентрацией действующего вещества в качестве ингибиторного субстрата¹¹.

Практическую проверку того, действительно ли изменение концентрации следует хилловской кинетике, наиболее наглядно можно провести, линеаризовав зависимость (9) путем использования подходящей комбинации переменных¹², обозначенных квадратными скобками в формуле ниже¹³:

$$[\log(m^{-1} - 1)] = -\alpha \log(C_{0.5}^\alpha) + \alpha [\log(C)], \quad (14)$$

где логарифм может быть взят по любому основанию.

Рис. 4 демонстрирует соответствие исследуемых данных кинетике Хилла, так как найденные медианы на переходном участке (точки, соответствующие медианным концентрациям равным или меньшим обоих контролей, исключены как приводящие к расхождению или мнимым значениям логарифма) в координатном представлении (14) с логарифмами по основанию два в силу двоичной последовательности разведений, с хорошей точностью аппроксимируются прямой. Тангенс ее угла наклона определяет степенной коэффициент $\alpha = 2.33$, а сдвиг – концентрацию половинного эффекта, близкую к 0.16 mg/ml.



¹⁰ J. Zhi, C. H. Nightingale, R. Quintiliani, A pharmacodynamic model for the activity of antibiotics against microorganisms under nonsaturable conditions, *Journal of Pharmaceutical Sciences* 75 (1986) 1063–1067. doi:10.1002/jps.2600751108

¹¹ S. Goutelle, M. Maurin, F. Rougier, X. Barbaut, L. Bourguignon, M. Ducher, P. Maire, The Hill equation: a review of its capabilities in pharmacological modelling, *Fundamental & Clinical Pharmacology* 22 (2008) 633–648. doi:10.1111/j.1472-8206.2008.00633.x; and R. Gesztelyi, J. Zsuga, A. Kemeny-Beke, B. Varga, B. Juhasz, A. Tosaki, The Hill equation and the origin of quantitative pharmacology, *Archive for History of Exact Sciences* 66 (2012) 427–438. doi:10.1007/s00407-012-0098-5

¹² По шагам: записав (9) в виде

$$m(C) = \frac{1^\alpha}{1 + \frac{C^{\frac{\alpha}{2}}}{C_{0.5}^{\frac{\alpha}{2}}}},$$

перевернув, $1 + C^{\frac{\alpha}{2}}/C_{0.5}^{\frac{\alpha}{2}} = 1/m$, перенес единицу в правую часть и взяв логарифм от обеих частей.

¹³ R. Barlow, J. F. Blake, Hill coefficients and the logistic equation, *Trends in Pharmacological Sciences* 10 (1989) 440–441. doi:10.1016/S0165-6147(89)80006-9

Рис. 4: Зависимость нормированных безразмерных медиан флуоресценции (маркеры) от концентрации действующего вещества в логарифмических координатах и фитующая их прямая линия, что свидетельствует о соответствии отклика кинетике Хилла.

Практическая аппроксимация

Так как ряд действующих веществ обладают собственными флуоресцентными свойствами, что приводит к сдвигу регистрируемого стационарного значения при высоких концентрациях от величины 1% контроля, с практической точки зрения удобнее проводить обезразмеривание не как в идеальном случае (7), а просто на величину нормального контроля¹⁴

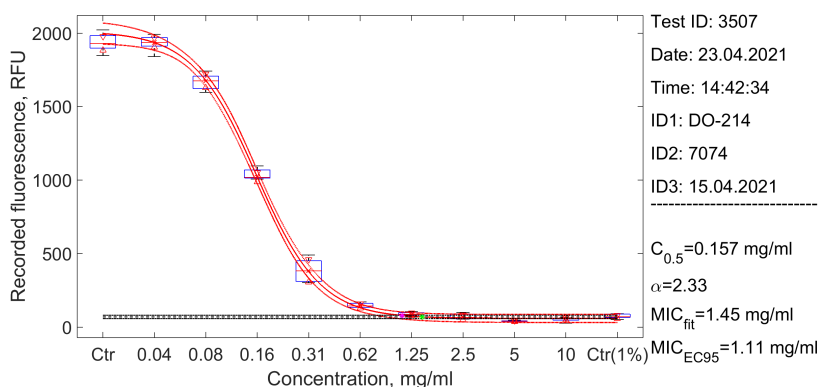
$$m = \frac{M}{M_{contr}} \quad (15)$$

С практической точки зрения, гиперболический тангенс (8) можно заменить эквивалентной формой логистической кривой, явно выраженной через экспоненту, как это было показано выше и, вдобавок, учесть в виде явных неопределенных параметров верхнее m_{max} и нижнее m_{min} значения ее асимптот, что соответствует стандартной методике фиттинга данных зависимости отклика на действующую дозу¹⁵:

$$m(\ln(C)) = m_{min} + \frac{m_{max} - m_{min}}{1 + e^{\alpha(\log(C) - \log(C_{0.5}))}} \quad (16)$$

Использование логарифма концентрации несмотря на наличие явной формулы зависимости от концентрации как таковой, (9), связано с тем, что гиперболическая функция (9) является быстро убывающей, а последовательность экспериментальных значений концентрации изменяется равномерно именно в логарифмическом масштабе¹⁶.

В качестве МИК для непрерывной аппроксимации используется концентрация, при которой $m = M_1/M_{contr}$; кроме того, на рис. 5 указаны доверительные интервалы аппроксимации, что, в принципе, позволяет оценить интервал неопределенности МИК по их пересечению с линией верхней границы доверительного интервала 1% контроля при росте концентрации.



¹⁴ Вдобавок, это позволяет работать с ситуацией записи флуоресценции без разбавленной культуры как таковой. Кроме того, в случае малого роста для 1% контроля в программе реализуется условие использования в его качестве 0.1 значений для нормального роста.

¹⁵ H. Motulsky, A. Christopoulos, Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting, Oxford University Press, 2004

¹⁶ См. комментарии по поводу устойчивости алгоритма аппроксимации в книге¹⁵.

Рис. 5: Экспериментальные данные по флуоресценции, построенные как диаграмма размаха, аналогичные рис. 2; красными ×-ами помечены точки на линиях, обозначающих медиану, использованные при фиттинге. Сплошная красная линия – фит функцией (16), штриховые – 95% доверительный интервал аппроксимации; зеленой и фиолетовой звездочками обозначены положения МИК для равенства кривой отклика и медианы контроля и 5% превышения величины нижнего стационара кривой отклика, соответственно.

Помимо этого, указываются также свойства отклика на действующее вещество, не требующие сравнения с величиной флуоресценции 1% контроля, а заданные свойствами самой аппроксимирующей кривой: (i) полумаксимальная эффективная концентрация $C_{0.5} \equiv EC_{50}$, соответствующая эффекту, равному половине возможного для данного вещества; (ii) MIC_{EC95} , равная концентрации, при которой функция отклика, аппроксимирующая данные, достигает $0.05m_{min}$, что рассматривается как критерий выхода на стационарное значение; при дальнейшем росте концентрации изменение флуоресценции не является существенным с учетом статистического разброса измерений и соответствует стандарту эффективной концентрации EC_{95} .

Исторические заметки

Методика, количественно использующая сравнение флуоресценции с уровнем контроля под воздействием антимикобактериальных соединений при резазуриновом тесте была предложена в 1997 году¹⁷ (с использованием критерия EC_{90} по уровню флуоресценции по сравнению с контрольной микобактериальной культурой). К 2007 году такой количественный метод уже рассматривался как стандартный протокол как с двоичными разведениями в рядах планшета, так и с более детальным разрешением по концентрации^{18,19}.

Использование фитирующей функции (16) является стандартным рекомендованным методом для исследований лекарственной чувствительности (*dose-response curve*), см. книгу¹⁵; в работах по лекарственной чувствительности микобактерий в резазуриновом тесте соответствующее сглаживание начинает активно использоваться в 2012–2015 гг.^{20,21}.

Более того, именно использование четырехпараметрической функции (16), рекомендовано на шаге обработки данных современных планшетных ридеров флуоресценции при резазуриновом тесте; более того, вывод соответствующих параметров (в том числе EC_{50}) включено в прилагающее программное обеспечение, причем именно и к тому ридеру, на котором анализируемые данные и снимались в НИИФ, см. стр. 48 инструкции:

- FLUOstar OPTIMA: Software Manual - Part IIIa: MARS Data Analysis

¹⁷ L. A. Collins, S. G. Franzblau, Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*, Antimicrobial Agents and Chemotherapy 41 (1997) 1004–1009. doi:10.1128/AAC.41.5.1004

¹⁸ N. K. Taneja, J. S. Tyagi, Resazurin reduction assays for screening of anti-tubercular compounds against dormant and actively growing *Mycobacterium tuberculosis*, *Mycobacterium bovis* BCG and *Mycobacterium smegmatis*, Journal of Antimicrobial Chemotherapy 60 (2) (2007) 288–293. doi:10.1093/jac/dkm207

¹⁹ T. P. Primm, S. G. Franzblau, Recent advances in methodologies for the discovery of antimycobacterial drugs, Current Bioactive Compounds 3 (2007) 201–208

²⁰ M. Zhang, C. Sala, R. C. Hartkoorn, N. Dhar, A. Mendoza-Losana, S. T. Cole, Streptomycin-starved mycobacterium tuberculosis 18b, a drug discovery tool for latent tuberculosis, Antimicrobial Agents and Chemotherapy 56 (2012) 5782–5789. doi:10.1128/AAC.01125-12

²¹ B. Gold, J. Roberts, Y. Ling, L. L. Quezada, J. Glasheen, E. Ballinger, S. Somersan-Karakaya, T. Warrior, J. D. Warren, C. Nathan, Rapid, semiquantitative assay to discriminate among compounds with activity against replicating or nonreplicating mycobacterium tuberculosis, Antimicrobial Agents and Chemotherapy 59 (2015) 6521–6538. doi:10.1128/AAC.00803-15