RESEARCH ARTICLE

Prediction of recombinant *Mycobacterium tuberculosis* α-crystallin oligomer chaperone activity using polynomial graphs [version 1; peer review: 2 approved]

Gautam Krishnan, Utpal Roy

Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Department of Biological Sciences, Zuari Nagar, Goa, 403726, India

**Abstract**

**Background:** Mycobacterial α-crystallin (Acr) is a chaperone that prevents misfolding of proteins when *Mycobacterium tuberculosis* is found in a latent form in the host tissue.

**Methods:** Using insulin as a model substrate and utilizing polynomial graphs, we attempted to predict molecular-level interactions that are a function of the oligomeric state of the recombinant protein. The chaperone activity of the recombinant oligomeric Acr was measured at 60°C with Acr samples obtained before gel filtration chromatography and compared with a gel-filtered sample.

**Results:** The polynomial graphs constructed showed improved molecular coverage of the insulin B chain by the oligomer. The 2nd order coefficient is the one that changes with the oligomeric ratio of Acr and improves chaperone activity. Polynomial analysis suggested that it could be a useful parameter to predict chaperone activity for potential *in vitro* batches of *M. tuberculosis* Acr based on the dynamic nature of the association and disassociation of oligomers.

**Conclusions:** The results showed that coverage of insulin B chain improved with higher ratio of 9-mer as compared to lower ratios. This was shown by both simulation plots and actual assay data. The polynomial graphs showed increase in the 2nd order coefficient, thus suggesting the important role of oligomerisation in improved molecular coverage of insulin B chain.

**Keywords**

Alpha-crystallin, Chaperone, Purification, Recombinant, Polynomial
Corresponding author: Utpal Roy (utpalroy010@gmail.com)

Author roles: Krishnan G: Conceptualization, Investigation, Methodology, Writing – Original Draft Preparation; Roy U: Data Curation, Formal Analysis, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2018 Krishnan G and Roy U. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

How to cite this article: Krishnan G and Roy U. Prediction of recombinant Mycobacterium tuberculosis α-crystallin oligomer chaperone activity using polynomial graphs [version 1; peer review: 2 approved] F1000Research 2018, 7:1801
https://doi.org/10.12688/f1000research.16328.1

First published: 15 Nov 2018, 7:1801 https://doi.org/10.12688/f1000research.16328.1
Introduction
Mycobacterial α-crystallin (Acr) is an important protein in latent tuberculosis (TB). When infected by *Mycobacterium tuberculosis*, the host survives the infection, yet in 1 out of 3 cases, the bacteria survive in a hypoxic state (Sherman et al., 2001). The alternative to *in vivo* studies is to clone and express recombinant Acr in *Escherichia coli* and study its chaperone activity *in vitro* (Gu et al., 2002; Panda et al., 2017; Yang et al., 1999). Substrates used for past enzymatic studies have included citrate synthase and insulin (Chang et al., 1996; Gu et al., 2002; Mao et al., 2001; Panda et al., 2017). However, there is no precise molecular means to quantify the biological activity of Acr in terms of molecular interactions. Acr is active as an oligomer and though its mechanism of action is known to be in the form of a trimer of trimers (Chang et al., 1996) or in the form of a dodecamer (Kenna-way et al., 2005; Panda et al., 2017), yet it is difficult to study the mechanism by which it inhibits the aggregation of substrates, either thermally or by dithiothreitol (DTT) induced with insulin.

In the present study, we analysed recombinant Acr preparations in terms of chaperone activity against the insulin B chain substrate, and explored the possibility of using the polynomial graphs in predicting chaperone activity. A theoretical and mechanistic representation using the real data has been used to render the numerical predictions of activity based on polynomial graphs.

Methods

Cloning and expression

The *acr* gene of *M. tuberculosis* H37Rv was directionally cloned using the expression vector pET28a as an N-terminal His tag into pET28a using NdeI and XhoI restriction enzymes, Takara. Expression of the recombinant acr was initially done at 50 ml scale, using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) induction (37°C for 3 h). The cell pellets obtained were used for the SDS-PAGE (15% acrylamide) to check for the expression of protein. The rest of the sample was sonicated with 10 mM Tris pH 7.0 / 5% glycerol followed by centrifugation at 20,000 g at 4°C for 30 min. Aliquots of supernatant and pellet obtained were used for the 15% SDS-PAGE gel to observe the localization of protein in supernatant or pellet.

Recombinant Acr was purified using Nickel-NTA agarose. The cell pellet obtained after 0.5–1.0 mM IPTG induction was lysed by sonication. The sonicate was centrifuged at 20,000g for 30 mins at 4°C and the supernatant obtained was allowed to bind to a 3 ml of Nickel-NTA resin. The protein was eluted with a 3-step gradient of 300 mM, 400 mM and 500 mM imidazole in buffer containing 20 mM Tris pH 7.0, 300 mM NaCl and 5% glycerol.

The Nickel-NTA eluted fractions containing the Acr were dialysed against 20mM Tris pH 7.0, 100 mM NaCl and 5% glycerol and subjected to gel filtration chromatography on a Sephacryl-200 Hiprep XK 16/60 Column (Pre-packed 120 ml) using the AKTAPurifier system (GE Healthcare Life Sciences). The column was equilibrated using the same buffer and calibration carried out using BIORAD standards. Molecular weight ranging from 670 kDa to 1.5 kDa. SDS–PAGE analysis was carried out with the first 3 lanes loaded with non-gel-filtered samples and the lanes 5–10 with 2 eluted peaks C1 and C2 of different runs with varying amounts of beta mercaptoethanol. A Native-PAGE analysis was carried out for both non-gel-filtered and gel-filtered samples, respectively using 8–16% gradient Tris glycerine gel, and bovine serum albumin (Hi-Media; 0.5 mg/ml) as a standard. The oligomer size was estimated using a plot of log molecular weight versus distance migrated and estimated sizes calculated using antilog.

Chaperone assay using insulin as substrate

Activity of recombinant *M. tuberculosis* Acr was assessed at 60°C using three concentrations (83, 118 and 125 μM) of insulin B chain as a substrate, by adding together Acr along with the insulin B chain. Aggregation of the insulin B chain was initiated by addition of 25 mM DTT and the change in the absorbance was monitored at 360 nm for 20 mins in kinetic mode with 1800 time points of 1 second interval. The assay was carried out at 60°C with His-tag-eluted samples at single point three concentrations of 12, 30 and 37.5 μM labelled as non-gel-filtered samples, and also with a His-tag plus gel-filtered sample at a concentration of 12 μM. The assay volume for 12 μM Acr was 400 microlitre while the other 2 assays were done in assay volume of 250 microlitres. Later assays were done at single-point concentrations of 11 μM and 37.5 μM at 60°C (data not shown).

Polynomial graph analysis

A set of simulation graphs was plotted in MS Excel 97-2003 assuming different ratios of oligomers and the number of molecules of Acr versus the number of molecules of insulin at 118 μM and the percentage of molecules of insulin B chain covered. The assumption was 4 molar concentrations of 5, 10, 20 and 40 μM and 2 different proportions of 9-mer to 12-mer oligomers; 60% of 9-mer and 10% each of 10- to 12-mer covered. The assumption was 4 molar concentrations of 5, 10, 20 and 40 μM and 2 different proportions of 9-mer to 12-mer oligomers; 60% of 9-mer and 10% each of 10- to 12-mer and 20% each of 9-mer to 12-mer. This gave an indication about the trends in percentage insulin B chain covered at two different ratios in scenarios, one with higher amount of nonamer and the second a lower amount of nonamer.

The assay data we obtained was used to calculate the percentage inhibition of insulin aggregation versus percentage of molecules of insulin B chain covered with both monomers and oligomers for non-gel-filtered sample. The assumption made was the molecular size of Acr to be 18 kDa as monomer and further plots were constructed assuming the proportion of 9-mers to 16-mers as obtained from Native-PAGE data. The polynomial plots were used to gain insights into the binding of Acr to the insulin B chain.

Results

Expression and activity studies

The His-tag purification and gel filtration showed greater than 95% purity as seen in the gel analysis along with upper (higher) molecular weight bands (Figure 1a). The protein appeared in the void volume suggested a blend of oligomers which was later confirmed by Native-PAGE analysis.
Native-PAGE analysis
Native-PAGE analysis showed a proportion of 70% of 9-mer and 10% each of 10-mer to 12-mer in the non-gel-filtered (before the gel-filtration chromatography) samples and 40% of 16-mer and 15% each of 9-mer to 12-mer as analysed by ImageJ software (Figure 1b, c and Table 1).

Chaperone activity at 60°C
The acr-pET28a construct showed activity against DTT induced aggregation of insulin. At a concentration of 37 μM of Acr and 125 μM insulin it gave 40% inhibition (Figure 2a). At a concentration of 12 μM Acr and 83 μM insulin it gave 60% inhibition (Figure 2b). A separate assay was carried out using both non-gel-filtered and gel-filtered samples at 12 and 30 μM, respectively and insulin at 118 μM concentration (Figure 2c). There was a difference in activity between both His-tag and gel-filtered samples. The non-gel-filtered samples showed 21.54% inhibition at a concentration of 30 μM, whereas 60% inhibition was achieved by the gel-filtered sample at a concentration of 12 μM Acr.

Polynomial graph analysis
Activity of Acr was plotted in terms of percentage inhibition versus percentage of molecules of insulin B chain and polynomial graphs constructed to check for best fit analysis with $R^2$ greater than 0.95 being the criteria (Figure 3a, b). The graphs showed 2nd order polynomials as the best fit. The coefficient of $x^2$ was higher in the oligomer plot, indicating that oligomerisation decreases the percentage of molecules of the insulin B chain covered. This was performed for the non-gel-filtered samples for both monomer and oligomer, assuming the ratios obtained from Native-PAGE analysis. The simulations showed the proportion of oligomers required to cover a maximum percentage of insulin B chain changed with the proportion of 9-mers with higher proportion of 9-mer covering more of the insulin B chain (Figure 4a, b).

Discussion
His-tag purification attempts of the recombinant Acr-producing construct showed the protein obtained was greater than 95% purity in the fractions eluted with 500 mM imidazole. The chaperone activity varied with two different concentrations of Acr, and the plots of monomer versus oligomer obtained showed that oligomerisation helps more molecular coverage the insulin B chain. This was confirmed by the results obtained by the gel-filtered samples that showed higher inhibition than the non-gel-filtered samples. This could be possibly explained by the higher proportion of 10-mer to 12-mer and 40% of 16-mer, which indicates that a higher proportion of oligomers improves chaperone activity.

The generation of polynomial graphs plotted showed an increase in the $x^2$ coefficient of the 2nd order graph, which is directly proportional to the amount of oligomers present and can be possibly used to model chaperone activity using insulin as a substrate. These observations suggest that the change in the percentage of molecules of insulin B chain covered is due to the extent of oligomerisation that forms and also varies with the...
Table 1. Oligomer ratio calculation. The Native-PAGE data was used to calculate the oligomeric sizes and ratios of different bands in non-gel-filtered and gel-filtered samples. Oligomer sizes were calculated using bovine serum albumin (BSA) monomers, dimers and trimers as a reference (66, 132 and 198 kDa) and the log of molecular weights plotted versus the distances travelled in centimetre. From this calculation, molecular weight was extrapolated for the different bands seen in non-gel-filtered samples (band 1, 2, 3, 4) and gel-filtered samples (band 1, 2, 3, 4 and 5).

| BSA (Mol. Wt.) | Log Mol wt. | Distance (cm) | Size(kDa) | Monomer(18 kDa) | Fraction of total |
|----------------|-------------|---------------|-----------|-----------------|------------------|
| 66             | 1.82        | 1             |           |                 |                  |
| 132            | 2.12        | 2             |           |                 |                  |
| 198            | 2.29        | 2.8           | 18        |                 |                  |
| Non-gel-filtered band 1 | 2.19 | 2.3 | 155 | 9 | 70 |
| Non-gel-filtered band 2 | 2.258 | 2.6 | 177.13 | 10 | 10 |
| Non-gel-filtered band 3 | 2.297 | 2.8 | 198 | 11 | 10 |
| Non-gel-filtered band 4 | 2.332 | 3 | 215 | 12 | 10 |
| Ratio of 10-mer to 12-mer | | | | | 30 |
| Gel-filtered band 1 | Same as H | | | | 9-mer 70 |
| Gel-filtered band 2 | | | | | 10 |
| Gel-filtered band 3 | | | | | 11 |
| Gel-filtered band 4 | | | | | 12 |
| Gel-filtered band 5 | 2.453 | 4 | 283.79 | 16 | 40 |
| Ratio of 10-mer to 16-mer | | | | | 85 |
| 9-mer | | | | | 15 |

Figure 2. Chaperone Activity of Acr at 60°C. (a) Insulin was assayed at a concentration of 125 µM with α-crystallin (Acr) at a concentration of 37.5 µM in an assay volume of 400 µl by adding Acr to prevent DTT induced aggregation of B chain. (b) Insulin was assayed at a concentration of 83 µM with Acr added at a concentration of 12 µM in an assay volume of 250 µl. (c) Insulin was assayed at a concentration of 118 µM and using two concentrations of non-gel-filtered Acr (30 µM) and gel-filtered Acr (12 µM), in an assay volume of 250 µl.

ratios of oligomers directly affecting the activity, suggesting that this is a significant parameter for checking in vitro Acr activity.

Polynomial graphs have been used to analyse activity of proteins involving oligomers (Yeow & Clayton, 2007). Many oligomer proteins function as high-molecular-weight aggregates and polymers, and the mechanism of heat-shock proteins is for that reason, poorly understood (Mattoo, 2014). This approach of using polynomials can shed insights into the molecular mechanism of oligomers binding to substrates.

Conclusions
Using polynomial graph analysis, we suggest a predictive tool using percentage inhibition versus insulin B chain covered for in vitro Acr preparations that could also be extrapolated for
Figure 3. Chaperone activity of Acr non gel filtered versus % of insulin B chain covered. (a) A polynomial plot was used to plot percentage inhibition of non-gel-filtered α-crystallin (Acr) monomers against percentage of molecules of the insulin B chain covered at 60°C. (b) A polynomial plot was used to plot percentage inhibition of non-gel-filtered Acr oligomer versus the percentage of molecules of insulin B chain covered at 60°C. (c) The percentage inhibition was replotted versus percentage of molecules of insulin B chain at 60°C by calculating the molecules of Acr based on the proportion of oligomers obtained from Native-PAGE analysis at 37°C.

Figure 4. Simulation plots of Acr non gel filtered with different ratios of oligomer. (a) A plot of simulated Acr oligomer assuming 9-mer ratio of 60% and 10% of 10- to 12-mer versus percentage of molecules of insulin B chain covered at 118 µM. (b) A plot of simulated Acr oligomer assuming a 9-mer to 12-mer ratio of 20% versus the percentage of molecules of insulin B chain covered at 118 µM.

**Grant information**
The author(s) declared that no grants were involved in supporting this work.

**Acknowledgements**
Authors acknowledge the Director and the Vice Chancellor of BITS Pilani for providing necessary financial assistance in the form of visiting fellowship and institute contingency. We wish to thank Dr. Santanu Datta of Bugworks, G.K.’s co-supervisor, who provided invaluable technical assistance.
Chang Z, Primm TP, Jakana J, et al.: *Mycobacterium tuberculosis* 16-kDa antigen (Hsp16.3) functions as an oligomeric structure in vitro to suppress thermal aggregation. J Biol Chem. 1996; 271(12): 7218–7223.

Gu L, Abulimiti A, Li W, et al.: Monodisperse Hsp16.3 nonamer exhibits dynamic dissociation and reassociation, with the nonamer dissociation prerequisite for chaperone-like activity. J Mol Biol. 2002; 319(2): 517–26.

Kennaway CK, Benesch JL, Gohlke U, et al.: Dodecameric structure of the small heat shock protein Acr1 from *Mycobacterium tuberculosis*. J Biol Chem. 2005; 280(39): 33419–33425.

Krishnan G, Roy U: Dataset 1 in: Prediction of Chaperone Activity of Recombinant *M.tuberculosis* Acr Oligomer Using Polynomial Graphs. F1000Research. 2018. http://www.doi.org/10.5256/f1000research.16328.d21935

Mao Q, Ke D, Feng X, et al.: Preheat treatment for *Mycobacterium tuberculosis* Hsp16.3: correlation between a structural phase change at 60 degrees C and a dramatic increase in chaperone-like activity. Biochem Biophys Res Commun. 2001; 284(4): 942–947.

Mattoo RUH: Mechanistic insights into cytosolic molecular chaperones in protein unfolding and disaggregation. Thesis. Originally published at: University of Lausanne Posted at the University of Lausanne Open Archive http://serval.unil.ch Document URN : urn:nbn:ch:serval-BIB_EA49B7E040F21. 2014.

Panda AK, Chakraborty A, Nandi SK, et al.: The C-terminal extension of *Mycobacterium tuberculosis* Hsp16.3 regulates its oligomerization, subunit exchange dynamics and chaperone function. FEBS J. 2017; 284(2): 277–300.

Sherman DR, Voskuil M, Schnappinger D, et al.: Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. Proc Natl Acad Sci U S A. 2001; 98(13): 7534–7539.

Yang H, Huang S, Dai H, et al.: *Mycobacterium tuberculosis* small heat shock protein Hsp16.3 exposes hydrophobic surfaces at mild conditions: conformational flexibility and molecular chaperone activity. Protein Sci. Cambridge University Press. Printed in the USA. 1999; 8(1): 174–179.

Yeow EK, Clayton AH: Enumeration of oligomerization states of membrane proteins in living cells by homo-FRET spectroscopy and microscopy: theory and application. Biophys J. 2007; 92(9): 3098–3104.
Open Peer Review

Current Peer Review Status: 🔄 ✅

Version 1

Reviewer Report 30 June 2020

https://doi.org/10.5256/f1000research.17837.r61501

© 2020 Ehtesham N. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nasreen Z. Ehtesham
National Institute of Pathology, Safdarjung Hospital, New Delhi, India

The manuscript titled “Prediction of recombinant *Mycobacterium tuberculosis* α-crystallin oligomer chaperone activity using polynomial graphs” by Gautam Krishnan and Utpal Roy demonstrated the relationship between various higher-order oligomers of Acr and its relation to the prevention of aggregation of insulin B chain as a function of its chaperone activity, predicted using a mathematical polynomial graph. Earlier reports showed that Acr exists in multiple higher-order oligomers, which is related to their chaperone functions. Acr is a vital chaperone protein expressed majorly in the stationary phase as well as in hypoxic conditions. It is required for the survival of *M. tb* in macrophages; thus, it is a critical virulence effector. The authors utilize the previously known pieces of information and use of polynomial graphs in the prediction of Acr oligomers and its relation to its chaperone activity. They devised a novel way to predict Acr chaperone activity and its relation to the oligomeric state of the Acr. The manuscript is well written, organized, and presented. Few minor concerns which are mentioned below should be addressed before finalizing the manuscript for indexing.

Introduction:
- The last sentence of the first paragraph needs revision. Minor grammatical error is present.

Results:
1. Acr of *Mycobacterium tuberculosis* exists in multiple oligomeric states, as reported earlier, and also mentioned by the authors in the introduction section. As shown in the denaturing gel picture, only prominent SDS resistant dimeric or trimeric forms are visible, other oligomers are present only in minute quantity. Is the dimeric form being the stable and primary oligomeric state of Acr in *M. tb*? The authors should also label gel pictures indicating Molecular weight standards and various oligomers.

2. How the authors calculated the percentage of various oligomers of Acr? Whether they used three gels run with samples purified in different batches or the results given are from a single gel?

3. The sentence starting under the heading “Chaperone activity at 60°C” needs revision. The
The manuscript titled “Prediction of recombinant Mycobacterium tuberculosis α-crystallin oligomer chaperone activity using polynomial graphs” by Gautam Krishnan and Utpal Roy demonstrated the relationship between various higher-order oligomers of Acr and its relation to the prevention of aggregation of insulin B chain as a function of its chaperone activity, predicted using a mathematical polynomial graph. Earlier reports showed that Acr exists in multiple higher-order oligomers, which is related to their chaperone functions. Acr is a vital chaperone protein expressed majorly in the stationary phase as well as in hypoxic conditions. It is required for the survival of M. tb in macrophages; thus, it is a critical virulence effector. The authors utilise the previously known pieces of information and use of polynomial graphs in the prediction of Acr oligomers and its relation to its chaperone activity. They devised a novel way to predict Acr chaperone activity and its relation to the oligomeric state of the Acr. The manuscript is well written, organized, and presented. Few minor concerns which are mentioned below should be addressed before finalizing the manuscript for indexing.

Responses from the authors

Introduction:
Acr is active as an oligomer and though its mechanism of action is known to be in the form of a trimer of trimers (Chang et al., 1996) or in the form of a dodecamer (Kennaway et al., 2005; Panda et al., 2017), yet it is difficult to study the mechanism by which it inhibits the aggregation of substrates, either thermally or by dithiothreitol (DTT) induced with insulin.  

**Author's response:** Suggested change: Acr is active in oligomeric state, and though it is known to act as a trimer of trimers (Chang et al., 1996) or a dodecamer (Kennaway et al., 2005; Panda et al., 2017), yet it is difficult to study the mechanism by which it inhibits the thermal aggregation or dithiothreitol (DTT)-induced aggregation of insulin.  

**Results:**

Acr of *Mycobacterium tuberculosis* exists in multiple oligomeric states, as reported earlier, and also mentioned by the authors in the introduction section. As shown in the denaturing gel picture, only prominent SDS resistant dimeric or trimeric forms are visible, other oligomers are present only in minute quantity. Is the dimeric form being the stable and primary oligomeric state of Acr in *M. tb*? The authors should also label gel pictures indicating Molecular weight standards and various oligomers.

**Author's response:** The dimer and trimer forms are of the Bovine serine albumin (BSA) standard in Lane 1 and 9 (Figure 1a). These are of the standard. The gel image has been relabelled. The predominant form of *M.tb* Acr is nonamer to 12-mers which could be an aggregation of trimers or dimers. Chang et al, 1996 showed that the repeating unit is a trimer and a trimer of trimers forms a nonamer. We need to do more experimentation to verify the same. The re-labelling of the images has been done as suggested by the Honorable reviewer. As can be seen in Figure lanes 3, 7 and 8 there is a proportion of 9 to 12 mers in the non-gel filtered samples as compared to higher proportion of 12 mers in lanes 4, 5 and 6.

How the authors calculated the percentage of various oligomers of Acr? Whether they used three gels run with samples purified in different batches or the results given are from a single gel?

**Author's Response:** The percentage of various oligomers was calculated using software called Image J. This was accomplished by measuring the intensity of each band calculating the total intensity and dividing the intensity of each band by the total intensity to obtain the proportion. The molecular weight of oligomers was obtained by comparing with the standard Bovine serum albumin standards. The results are from a single gel. All batches were analyzed similarly and showed a range of oligomers concentrations which could explain the relation between higher oligomer ratio and activity.

The sentence starting under the heading “Chaperone activity at 60°C” needs revision. The authors should revise the sentence.

**Author's Response:** According to the suggestion heading has been changed to ‘Chaperone activity of Recombinant Acr at higher temperature’

**Competing Interests:** No competing interests were disclosed.
The Acr protein is a chaperone that prevents misfolding of proteins in latent *M. tuberculosis*. Characterizing the precise molecular mechanism of action of Acr, however, remains a challenge. The authors, Krishan and Roy, have used a novel means of estimating the degree of oligomerization using polynomial graphs. The authors make two key observations: the coverage of the insulin B molecules is influenced by (a) the degree of Acr oligomerization and (b) the ratio of oligomers.

This work lays the groundwork for further experiments in understanding the chaperoning activity of Acr at a molecular level.

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
Yes

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Infectious disease, bioinformatics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com