JOURNAL OF
EMERGING INVESTIGATORS

Article

The characterization of quorum sensing trajectories of *Vibrio fischeri* using longitudinal data analytics

Maryam Abdel-Azim[1], Ahmad Abdel-Azim[2]
[1] Central Buck High School East, Doylestown, Pennsylvania
[2] Harvard University, Cambridge, Massachusetts

SUMMARY

Quorum sensing (QS) is the process in which bacteria recognize and respond to the surrounding cell density, and it can be inhibited by certain antimicrobial substances. In *Vibrio fischeri*, bioluminescence intensifies as more cells utilize QS, making it an accurate biomarker for QS. In this study, we showed that observed illumination intensity data measured at multiple time points is insufficient for evaluating and comparing QS activity across different cultures without proper statistical modeling. We introduced a statistical approach to characterize illumination trajectories over time using longitudinal data analytics. The approach was tested on the naturally illuminating bacteria, *V. fischeri*, which emit light when QS is activated, but it can also be implemented with green fluorescent protein (GFP) and its numerous applications. To compare multiple QS trajectories, we suppressed QS activity at varying levels using honey, known for its antimicrobial and anti-QS properties. We characterized and compared illumination trajectories through time in bacterial cultures containing five Manuka honey concentrations. The longitudinal approach analyzed illumination data throughout the entire QS cycle and correctly ranked bacterial cultures according to their true level of QS activity. Conversely, the conventional cross-sectional approach was inconsistent in evaluating bacteria for their QS activity and trajectory. We conclude that modeling illumination intensity through time provides a more accurate evaluation of QS activity than conventional cross-sectional analysis.

INTRODUCTION

Quorum sensing (QS) is the process of cell-to-cell communication for the purpose of acquiring information about cellular density in the surrounding environment (1–3). Cellular density enables bacteria to behave in a collective, synchronized manner to regulate gene expression (1). Cells communicate by releasing signaling molecules called autoinducers. Other cells detect autoinducers as they are released into the environment. Once enough autoinducers are recognized, bacteria then express certain traits, such as illumination in the Gram-negative marine organism *Vibrio fischeri* (1, 2). Over time, bioluminescence peaks, gradually fades, then eventually disappears through QS (3).

Many compounds have been reported to inhibit QS, such as quorum-quenching enzymes and honey (4–6). Quorum-quenching enzymes prevent QS by degrading the autoinducers released by each cell (4). One example is Manuka honey (MH), which shows anti-QS properties (7). MH can eliminate autoinducers in the environment while impacting the expression of certain genes involved in the QS cycle in bacteria (8, 9). Further, honey is an antimicrobial substance that slows down bacterial growth at low concentrations and delays the activation of QS (7–9).

To measure and compare QS activity in different cultures, other studies relied mainly on average or peak illumination (10, 11). However, modeling the full QS activity cycle is crucial in characterizing QS curves that often have different shapes, delayed or earlier peaks, and varied illumination levels. Examining only a certain characteristic or a cross-sectional time-point of the curve excludes other curve details, leading to inaccurate results, so in this study, we modeled the full illumination intensity trajectories to evaluate and correctly rank QS activity in bacteria.

To create different trajectories of illumination cycles, we variably inhibited QS, utilizing different concentrations of MH. To then characterize the QS cycles through time, we modeled QS trajectories by fitting Legendre polynomials within each honey concentration. We hypothesized that observing several curve characteristics through time can be used to provide a more accurate evaluation of QS activity in bacteria. Using cross-sectional or regular regression approaches to compare QS cycles was shown to be less accurate as compared to the proposed longitudinal approach. Longitudinal approaches can efficiently differentiate between the effects different substances, antibiotic drugs, and even genes have on the QS cycles of bacteria. The approaches used in this study can be implemented to not only *V. fischeri*, but also pathogenic bacteria. The robust approaches allow for illumination data to be accurately analyzed, thus providing information about the factors impacting QS in bacteria.

RESULTS

We first sought to examine and model multiple QS cycles of *V. fischeri* cultures varying in inhibitor strengths. To model and evaluate multiple QS (illumination) cycles, we used honey at concentrations ranging from 15–90 mg/mL as an anti-QS substance and a control treatment of 0 mg/mL of honey

DOI: https://doi.org/10.59720/22-227

(**Figure 1**). This created QS cycles with known trajectories; for example, the culture containing 15 mg/mL MH concentration (MH15) was expected to result in higher QS activity than 30 mg/mL MH (MH30). We periodically recorded longitudinal illumination data and modeled using 7th-degree Legendre polynomials to estimate a curve for each concentration. A unique and accurate trajectory of illumination per MH concentration culture through time can be estimated by fitting separate time polynomials within each concentration. We then used these estimated curves to evaluate and rank the QS activity associated with cultures containing each honey concentration. Here, we used varying concentrations of honey to test the ability of the proposed longitudinal approach to rank QS activity in a consistent manner (i.e., greater honey concentrations result in lower QS activity).

Observed illumination data (**Figure 2**) were regressed on concentrations (0 to 90 mg/mL) to test the hypothesis of nonzero pairwise differences between concentrations. Differentiating between the QS cycles of the *V. fischeri* cultures with different treatments is difficult since the cycle rankings vary through time. We ranked the model-estimated illumination intensity of each concentration culture according to their expected average illumination intensity (**Figure 3A**). There were no significant differences between concentration pairs {MH30, MH60} and {MH60, MH90} (p>0.05, **Figure 3B**). Additionally, the QS cycles of treatment pairs {MH15, MH30} and {MH30, MH90} were significant at the 0.05 level.

Using partial data of the first 300 minutes limited the efficacy in evaluating QS activity (**Figure 4**). Regressing early illumination intensity data on concentration, while ignoring the rest of the QS cycle, failed to identify significant differences between concentration pairs {control, MH15}, {MH60, MH90} (p>0.05). Additionally, the pair {MH30, MH60} was significantly different at the 0.05 level.

To better characterize the change in illumination through time, we modeled QS trajectories within the five treatments studied by fitting a Legendre polynomial to the seventh degree within each concentration and the control treatment. First, a value for illumination intensity was estimated at each minute of the experiment, allowing for the development of a smooth curve within each treatment (**Figure 5**). Unlike the raw data of **Figure 2**, the properties of the cycle within each concentration were successfully differentiated when using the proposed curve approach (**Figure 6**). In addition to the clearly differentiated curves, we fitted the fitted curve data estimated at each minute as a dependent variable against concentrations as an independent variable in a regression model. All pairwise concentrations were highly significant (p<0.01, **Figure 6**).

In addition to correctly ranking QS activities, the longitudinal approach identified important characteristics associated with each treatment. Higher concentrations of honey not only depressed peak illumination, but they also delayed the activation of QS (**Figure 5**). Compared against the control treatment, all MH-treated cultures yielded shorter and weaker illumination cycles (**Figure 5**). The MH90-treated culture had the lowest and most delayed peak of illumination. In addition to average illumination, these curve features can be used to further evaluate the degree of QS activity in bacterial cultures. Greater MH concentrations resulted in

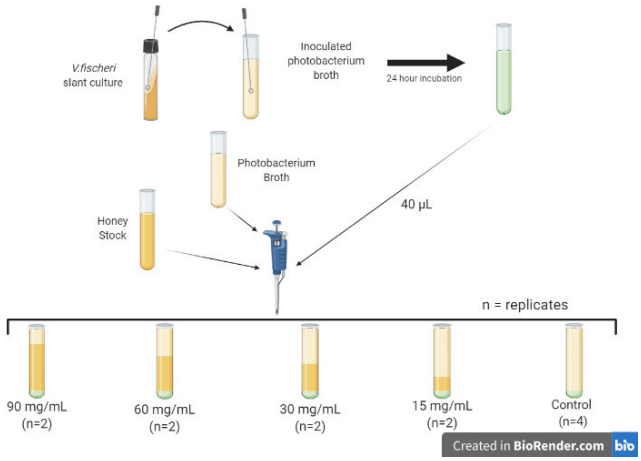

Figure 1: Experimental design to prepare cultures exposed to varying honey treatments. *V. fischeri* slant culture was used to prepare a liquid culture for the purpose of culturing five other vials containing varying amounts of photobacterium broth and honey. The same amount of bacteria was in each culture and the effect of honey was quantified throughout the study. Each MH concentration had 2 replicates. This figure was created using BioRender.com.

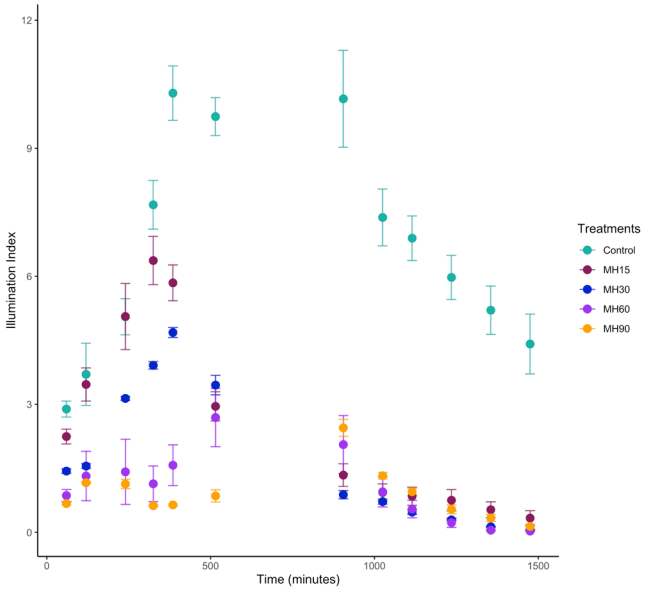

Figure 2: The ranking of raw illumination intensity within each concentration varies through time, thus making it difficult to differentiate between treatment efficacy. Quorum sensing cycles of V. fischeri with different MH concentrations (0, 15, 30, 60, and 90 mg/mL of honey concentrations) were obtained using raw illumination data collected over the course of 1440 minutes. Data points represent the average ± standard deviation illumination from each treatment replication (n=4).

DOI: https://doi.org/10.59720/22-227

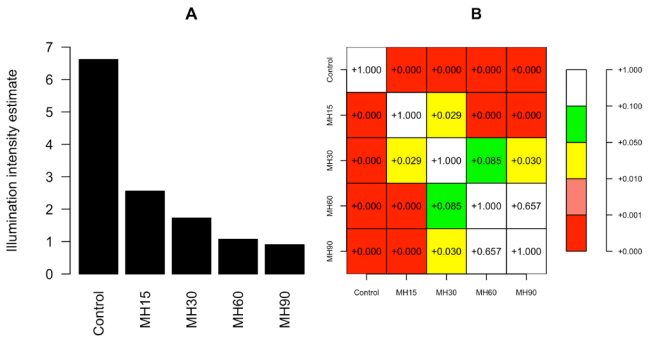

Figure 3: Accuracy of detecting and differentiating between different QS cycles was decreased when studying the observed illumination data. Eight concentration comparisons of observed QS activity data are significantly different from each other (p<0.05) with 6 comparisons being highly significant (p<0.01). Illumination intensity estimates for each concentration obtained from a regression model on all observed data are shown in (A). In the contrasts matrix (B), each square contains a p-value to test the difference between 2 concentrations. A p-value higher than 0.05 indicates a difference that is not significantly different from 0.

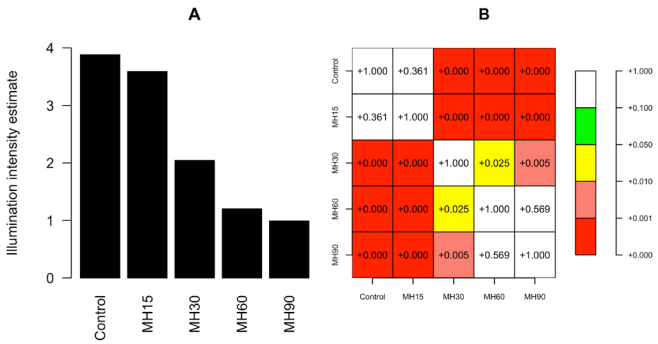

Figure 4: Cross-sectional analysis decreased accuracy of detecting and differentiating between different QS cycles. Eight concentration comparisons of observed QS activity data are significantly different from each other (p<0.05). (A) Illumination intensity estimates obtained from a regression model on observed data during the first 300 minutes of the QS cycle for each concentration. In the contrasts matrix (B), each square contains a p-value to test the difference between two MH concentrations (α = 0.05).

cultures having a more delayed and depressed QS cycles. We found that using cross-sectional regression approaches to compare QS cycles is less accurate as compared to the proposed longitudinal approach, which can efficiently and precisely differentiate between various levels of stress honey places on the QS cycles of bacteria.

DISCUSSION

To characterize the trajectories of QS in *V. fischeri* treated with five different inhibition concentrations and to study the change through time, we used longitudinal data approaches. While these approaches were able to detect and differentiate between QS cycles, conventional regression models failed since they ignored several important aspects of the QS cycle. For example, regular regression models revealed no significant differences between MH60 and either of MH30 or MH90 concentrations, while the longitudinal approach did. This study demonstrates the effectiveness and importance of accounting for several curve characteristics over time as the proposed longitudinal statistical approach yields more accurate results than regular cross-sectional approaches.

The illumination curves created using Legendre polynomials allowed for an accurate characterization of all concentration cycles. The difference between trajectories and their peaks was visible and identifiable because there was an estimated illumination value at every minute of the QS

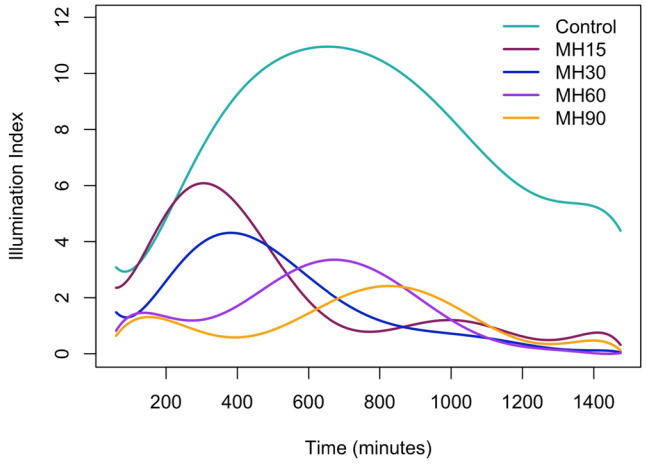

Figure 5: Ranking for treatment efficacy is clear when curve parameters are estimated within each treatment. Parameters estimated within concentrations of 0, 15, 30, 60, and 90 mg/mL were utilized to obtain an illumination estimate at every minute of the study (n=4). Illumination estimates represent the full QS cycle of each MH concentration treatment.

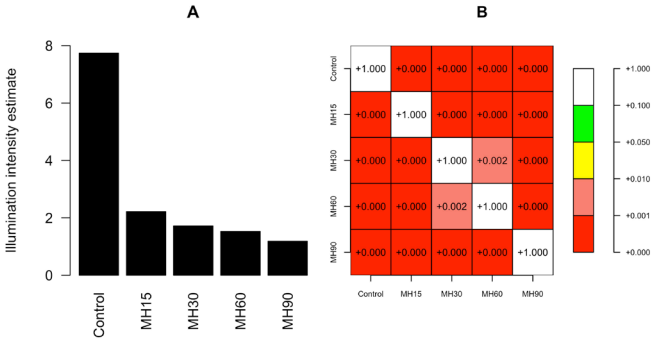

Figure 6: The proposed longitudinal approach yielded more accurate results than cross-sectional approaches. Ten concentration comparisons of observed QS activity data are highly significant (p<0.01). (A) Illumination intensity estimates for each concentration obtained from a regression model on longitudinally-modeled data . In the contrasts matrix (B), each square contains a p-value to test the difference between 2 concentration treatments (α = 0.05).

DOI: https://doi.org/10.59720/22-227

Order (p)	R-squared	Adjusted R-squared
2	0.9471	0.9442
5	0.9819	0.9798
6	0.9834	0.9812
7	0.9851	0.9827
8	0.9852	0.9825

Table 1: Polynomials of order 7 had the greatest adjusted R-squared value. R-squared and adjusted R-squared values vary for models with Legendre polynomials of orders 2 to 8. Polynomials of order 7 were utilized in the study due to its greater adjusted R-squared value.

cycle. It is necessary to characterize the entire cycle of each concentration and not to rely on select sections of the data that were insufficient in differentiating treatment effects.

Other studies, which rely only on the detection of illumination during certain parts of the QS cycle, might be missing important biological findings that contribute to the changes in illumination curves through time (10–13). Longitudinal data analytics are important in acquiring valuable characteristics by accurately modeling time trends. By measuring QS activity through time, we were able to discern the full effect of honey as an anti-QS substance on *V. fischeri*. Not only did it depress the entire QS cycle, but it also delayed the time of QS activation. This latter finding suggests that honey could be disrupting the ability of bacteria to sense its surroundings, so a higher cell density might be required for the bacteria to illuminate (8, 9). Since honey also affects bacterial growth, increased honey concentration treatments could not only be inferring with QS but also inhibiting cell proliferation. In further experiments, to differentiate the two processes, we could measure the absorbance of each vial to obtain a value for cellular density or use a hemocytometer to calculate the ratio of live to dead cells. Although the image analysis process was automated, the process of taking those digital images was not. Therefore, there are time points missing between minutes 500–850, so future experiments could avoid this problem by using a plate reader or an automated camera shutter system.

Longitudinal approaches can play a vital role in evaluating the efficacy of different substances or antibiotic drugs on QS activity in virulent bacteria. They also serve to evaluate the effect of a gene on QS activity in bacteria and contribute to a system's pathway analysis process. By using the approach proposed in this study, longitudinal illumination data can be analyzed to obtain more robust results about the factors that impact QS in bacteria.

MATERIALS AND METHODS

Broth preparation

Photobacterium broth (Carolina Biological) was prepared at a concentration of 65.8 g/L. The mixture was thoroughly mixed and sterilized in a pressure autoclave for 50 minutes. Honey stock of 300 mg/mL was prepared with hydrated photobacterium broth and was used to make 4 different media with honey concentrations of 15, 30, 60, and 90 mg/mL.

Bacterial culture

Sterilized broth was inoculated with *V. fischeri* bacteria (slant culture, Carolina Biological). The inoculated test tube was incubated for 24 hours on an orbital shaker set to 200 rpm at 25°C. This liquid bacterial culture was prepared and incubated at 25°C for the next steps of the experiment.

Variation in QS activity was created by using different concentrations of honey in bacterial cultures.

Two replicates of concentrations 15, 30, 60, and 90 mg/mL, along with four replicates of the control treatment of 0 mg/mL of the honey concentrations were used. Different amounts of media from the prepared concentrations were pipetted into their corresponding vials (**Figure 1**).

Acquisition of digital images

High-resolution images were taken every 1–2 hours over a period of 24 hours of every replicate in each concentration (Canon EOS Rebel T7i camera) after the introduction of honey. The digital images were taken inside a dark chamber with a 30 s camera exposure to capture the light emitted by the bacteria. The camera settings and position were fixed throughout the experiment for consistency. Each image contained six vials, and two shots were taken of each set, for a total of four shots at every time point.

Illumination detection from digital images

To obtain the average RGB value of each vial, images were cropped to sections with one vial each. RGB values of each pixel were averaged over an illuminating area and used to evaluate intensity of the area (3). The process was automated using Python and R studio to handle the hundreds of images taken throughout the experiment (14).

Proposed longitudinal approach

Average pixel intensities associated with the digital images of each vial were fitted as a dependent variable against treatments in a statistical model. In regular linear regression, the following model was used to estimate the effect of each concentration on illumination,

$$\boldsymbol{y} = X\boldsymbol{b} + \boldsymbol{e} \qquad [1]$$

where $\boldsymbol{y}$ is average pixel intensities, $\boldsymbol{b}$ is a vector of length 5 representing concentration effects on illumination, and $\boldsymbol{e}$ is a random residual. The matrix X is an incidence matrix of 288 rows by 5 columns of 0's and 1's. Rows of X represented observations, and columns represented the control, 15, 30, 60, and 90 mg/mL honey concentrations. For each observation, a 1 was placed in the column that corresponds to the concentration where the observation came from. All other elements in the row were labeled as 0's (**Figure 7A**).

To solve for the components of $\boldsymbol{b}$, we used equation [2] to obtain the regression coefficients of concentration on illumination:

DOI: https://doi.org/10.59720/22-227

$$\hat{\mathbf{b}} = (X^T X)^{-1} X^T \mathbf{y} \quad [2]$$

Note that $\hat{\mathbf{b}}$ is an estimate of the true $\boldsymbol{b}$ and X^T is the transpose of X.

Since the general time trend of illumination was being characterized, a polynomial of order p for time was fitted in the model to estimate $p+1$ parameters needed to characterize the general trend of illumination across all concentrations. The same model [1] can be used, except that the matrix X is modified to include the trend of time (**Figure 7B**). This, however, models a general trend through time across all treatments, but it does not achieve our objective of modeling a unique illumination trend specific to each concentration.

Instead of estimating a single effect per concentration or a general time trend of illumination, the objective was to characterize an illumination curve per concentration through time. Model [1] can again be modified by using the model matrix described in **Figure 7C**. It was shown that the estimated cycle is more accurate in distinguishing illumination curves than a single summary statistic (i.e., concentration effects).

It is more favorable to use the orthogonal Legendre polynomials because the covariance between them is minimal compared to fitting direct polynomials of time. Legendre polynomials were utilized here due to their simplicity compared to other orthogonal polynomials. In **Figure 7C**, each t_i is a random variable, of the scaled time in minutes, corresponding with observation i. Time is scaled from -1 to 1 using equation [3]:

$$t_i = -1 + 2(m_i - a)/(b - a) \quad [3]$$

where m_i is time in minutes of observation i, and a and b are the minimum and maximum times of the study. Note that in the model matrix of **Figure 7C**, polynomial coefficients replaced the 1's of the base model incidence matrix, expanding them to $(p+1)$ coefficients, where $p = 7$.

In this study, a polynomial of order 7 was chosen as it achieved the highest prediction accuracy as confirmed by the adjusted R-squared values (**Table 1**).

To further explain the statistical adjustment depicted in **Figure 7**, note that matrix X in **Figure 7A** corresponds to a model that attempts to estimate the effect of concentration without fitting time in a single-variable regression model. The structure of X in **Figure 7B** simply extends the model to multiple regression, where prediction is now based on multiple variables, concentration plus simple time polynomials. Finally, the structure in **Figure 7C** corresponds to another multiple regression where polynomials are now orthogonal instead of direct and are fit within each one of the concentrations studied. Fitting separate time polynomials within concentrations makes it possible to estimate a unique trajectory of illumination per concentration.

Regression analysis to test pairwise concentration differences

A linear regression model with illumination intensity as an independent variable and concentration as a dependent variable was fitted in R (4). The lm function in base R was run five times with the intercept removed and each concentration set as the reference level once. This allowed for the pairwise testing of zero differences between each two concentrations. We then constructed a matrix of contrasts to identify significant differences between each two concentrations for full and partial observed data, as well as for data predicted using the proposed longitudinal approach. A p-value less than 0.05 indicated that the difference between two varying concentration cultures was statistically significant.

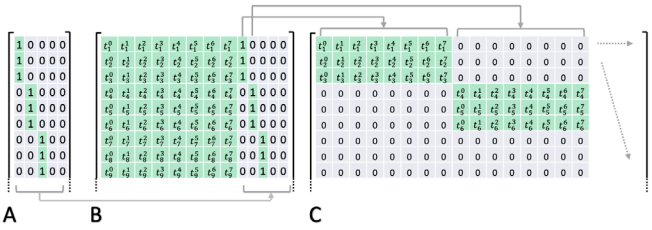

Figure 7: Example model matrices of 3 different models. (A) This is a regression model with concentration as the independent variable where the effect of each concentration level is estimated. The green blocks highlight concentration levels. (B) This is a regression model with two independent variables: polynomial for time trend and concentration. (C) Instead of a single time trend as in (B), here there is a time trend for each concentration level. The polynomials shown in (B) and (C) are Legendre polynomials. The superscripts indicate the polynomial order of time, and the subscripts indicate observation number.

Received: September 10, 2022
Accepted: October 2, 2022
Published: December 16, 2023

REFERENCES

1. Miller, Melissa B., and Bonnie L. Bassler. "Quorum Sensing in Bacteria." *Annual Review of Microbiology*, vol. 55, no. 1, 2001, pp. 165-199. https://doi:10.1146/annurev.micro.55.1.165.
2. Fuqua, Clay, *et al*. "Census and Consensus in Bacterial Ecosystems: The LuxR-LuxI Family of Quorum-Sensing Transcriptional Regulators." *Annual Review of Microbiology*, vol. 50, 1996, pp. 727-751. https://doi:10.1146/annurev.micro.50.1.727.
3. Abdel-Azim, Maryam G., and Gamal A. Abdel-Azim. "Characterizing Quorum Sensing-Induced Bioluminescence in Variable Volumes with Vibrio fischeri Using Computer Processing Methods." *Journal of Emerging Investigators*, vol. 2, 2020. https://doi.org/10.59720/20-010.
4. Rutherford, Steven T., and Bonnie L. Bassler. "Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control." *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 11, 2012, a012427. https://doi:10.1101/cshperspect.a012427.
5. McCann, Jessica, *et al*. "Population Dynamics of Vibrio fischeri during Infection of Euprymna scolopes." *Applied and Environmental Microbiology*, vol. 69, no. 10, 2003,

DOI: https://doi.org/10.59720/22-227

pp. 5928-5934. https://doi:10.1128/aem.69.10.5928-5934.2003.
6. Rémy, Benjamin, *et al*. "Interference in Bacterial Quorum Sensing: A Biopharmaceutical Perspective." *Frontiers in Pharmacology*, vol. 9, 2018, article 203. https://doi:10.3389/fphar.2018.00203.
7. Abdel-Azim, Salma G., *et al*. "Characterization of the Gain and Loss of Resistance to Antibiotics versus Tolerance to Honey as an Antimutagenic and Antimicrobial Medium in Extended-Time Serial Transfer Experiments." *Pharmacognosy Research*, vol. 11, 2019, pp. 147-154. https://doi:10.4103/pr.pr_175_18.
8. Wang, Rui, *et al*. "Honey's Ability to Counter Bacterial Infections Arises from Both Bactericidal Compounds and QS Inhibition." *Frontiers in Microbiology*, vol. 3, 2012, article 144. https://doi:10.3389/fmicb.2012.00144.
9. Prateeksha, *et al*. "Scaffold of Selenium Nanovectors and Honey Phytochemicals for Inhibition of Pseudomonas aeruginosa Quorum Sensing and Biofilm Formation." *Frontiers in Cellular and Infection Microbiology*, vol. 7, 2017, article 93. https://doi:10.3389/fcimb.2017.00093.
10. Pérez, Pablo Delfino, and Stephen J. Hagen. "Heterogeneous Response to a Quorum-Sensing Signal in the Luminescence of Individual Vibrio fischeri." *PLoS One*, vol. 5, no. 11, 2010, article e15473. https://doi:10.1371/journal.pone.0015473.
11. Abbas, Mazhar, *et al*. "Vibrio fischeri Bioluminescence Inhibition Assay for Ecotoxicity Assessment: A Review." *The Science of the Total Environment*, vol. 626, 2018, pp. 1295-1309. https://doi:10.1016/j.scitotenv.2018.01.066.
12. Abdel-Azim, Maryam. "Illumination Phenotyping Using Longitudinal Data Analytics (IPLA)." GitHub, 2022, github.com/maryamazim/.
13. Andersen, J. B., *et al*. "gfp-based N-acyl homoserine-lactone sensor systems for detection of bacterial communication." *Applied and Environmental Microbiology*, vol. 67, no. 2, 2001, pp. 575-585. https://doi:10.1128/AEM.67.2.575-585.2001.
14. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2021. https://www.R-project.org/.