Increasing the Production of dsRNA for Use as Biological Pesticides

Authors:

Allison Boboltz

Evan Biedermann

Joseph Frank

Words: 4885

Number of Figures: 8

Number of Tables: 0

Number of Equations: 0

Number of Supplements: 11

Number of References: 11

Increasing the Production of dsRNA for Use as Biological Pesticides

Allison Boboltz[a,1], Evan Biedermann[a,2], Joseph Frank[b]

[a] Biomedical Engineering, University of Virginia
[b] AgroSpheres, Inc.
[1] amb4mb@virginia.edu
[2] ecb3qb@virginia.edu

Abstract

The amount of chemical pesticide use is increasing as the agricultural industry expands to meet the growing global demand for food security. The increased application of pesticides poses a threat to human health, since studies have found that exposure to chemical pesticides can cause a number of physiological issues. AgroSpheres is a Charlottesville-based startup focused on developing biopesticides, a safer alternative to chemical pesticides. The aim of this capstone project is to advance the goal of phasing out chemical pesticides by improving biopesticide production at AgroSpheres. To achieve this, we designed and tested an exponential feed control algorithm to improve the company's fermentation process, and we characterized and optimized the expression conditions of the GFP-fused double-stranded RNA binding protein to increase the stability and longevity of the biopesticide. The results from testing the new feed algorithm found evidence that the new feeding method will likely be more effective and efficient than the previous method with further optimization to the design of the algorithm. The results of the protein study revealed clear expression patterns under various inducer concentrations and induction timing windows. The team faced limitations from the COVID-19 pandemic as well as AgroSpheres' production schedule which affected the ability to gather sufficient data. Despite the setbacks, our team was able to make improvements to AgroSpheres production process which the company can investigate further to help achieve the goal of providing a safer alternative to chemical pesticides.

Keywords: AgroSpheres, biopesticide, fermentation, dsRNA, protein expression,

Introduction

Chemical pesticides have been widely used for many years to protect agricultural crops from damage. Approximately 5.6 billion pounds of pesticides are used globally each year, with 1 billion pounds used in the United States alone (Alavanja, 2009). As the world population continues to grow, agriculture will also have to be scaled up to feed an increasing number of people, escalating pesticide use even further. This staggering annual pesticide use has caused a myriad of problems associated with human health and the environment. In particular, research has linked exposure to chemical pesticides to an increased risk of cancer and health issues within the gastrointestinal, reproductive, endocrine, neurological, respiratory, and reproductive systems (Mnif et al., 2011; Nicolopoulou-Stamati et al., 2016; Sanborn et al., 2007). People can be exposed to chemical pesticides through inhalation, skin contact, or ingestion (Nicolopoulou-Stamati et al., 2016). Residues on crops, in water, and in the air all contribute to the accumulation of pesticides within the human body and can cause negative effects (Nicolopoulou- Stamati et al., 2016). However, RNA interference (RNAi) biopesticides are a promising alternative to chemical pesticides.

RNAi is a highly effective method of targeting and eliminating pests, while also reducing human health risks (Mamta & Rajam, 2017). Chemical pesticides are harmful to human health because they target pest enzymes which can be homologous to human enzymes (Coman et al., 2013). However, RNAi biopesticides utilize double stranded RNA (dsRNA) sequences that are highly specific to the target organism. Once ingested by a pest, the dsRNA works to silence essential genes and prevent translation of mRNA into proteins (Kim & Rossi, 2008). Despite the advantages of RNAi biopesticides, dsRNA tends to degrade very rapidly, especially when applied topically to crops (Fletcher et al., 2020). This is the biggest barrier to

commercialization and so far has prevented the successful application of RNAi technology to agriculture. AgroSpheres is a company working to develop the first commercially viable RNAi biopesticide. The company has patented the use of bacterial minicells to encapsulate dsRNA (see Supplemental Figure 1).

Agrospheres uses a strain of *E. Coli* bacteria that produce target specific dsRNA sequences, the GFP-fused dsRNA binding protein, and minicells. The minicells bud off from the parent cells and contain the product dsRNA and GFP-fused dsRNA binding protein. The minicell is able to protect the dsRNA from rapid degradation when applied in the field. The GFP-fused dsRNA binding protein is another measure to prevent degradation by binding to the dsRNA to stabilize it. Minicells are also generally safe for human consumption because they are bioparticles that contain no chromosomes and are unable to proliferate (Islam, et al., 2021). The innovative use of minicell encapsulation promises to advance the development of RNAi biopesticides. However, certain elements of AgroSpheres' platform technology are currently insufficient for commercial production and adoption in the agriculture industry. Therefore, the objective of this capstone project is to enhance the efficacy of AgroSpheres' RNAi biopesticide through improving both the fermentation process and GFP-fused dsRNA binding protein expression.

Specifically, this capstone project aimed to improve the fed-batch fermentation process by implementing an exponential feed rate algorithm, which had never been previously attempted at the company. The fermentation bioprocess used to manufacture AgroSpheres' RNAi biopesticide must be able to successfully scale up to match production needs. Optimizing a fed-batch fermentation reaction is a process of balancing essential environmental factors to promote an optimal cell metabolism for growth and production. One of the most influential environmental factors is substrate concentration. If a substrate concentration is too low, the cells will not produce enough energy to grow and therefore will start to die. If the substrate concentration is too high, the cells will enter into overflow metabolism which will produce growth inhibiting metabolites such as acetate (Basan et al., 2015). Therefore, it is essential to design an optimal feeding strategy which will provide just enough nutrients to induce healthy, exponential growth. The predicted outcomes of utilizing an exponential feed rate were an increase in the production of dsRNA, cell density, and cell viability.

Another goal of this capstone project was to increase the stability and longevity of AgroSpheres' product. The second aim was to characterize GFP-fused double stranded RNA binding protein expression and optimize expression conditions in the soluble fraction, which had also never been done before by the company. Increasing expression in the soluble protein fraction is important because the product dsRNA being produced is soluble. In order to bind and stabilize the product dsRNA, we focused specifically on studying expression of the GFP-fused dsRNA binding protein in the soluble protein fraction. The team investigated the effects of changing inducer concentrations in two different media types and altering induction timing. The predicted outcomes of this were identifying the best expression conditions to maximize soluble GFP-fused dsRNA binding protein expression.

Materials and Methods

Feed Media and Feed Solution Preparation

All bioreactor growth runs were started with inoculation from a seed culture. All media recipes and components were provided from AgroSpheres standard operating procedures (SOPs). Media for a seed culture were prepared by combining Citrate Phosphate Buffer, 70 w/v % Glycerol, 2M $MgSO_4$, 100X Trace Metal Solution, 5 mg/mL Thiamine, 50 mg/mL Kanamycin, and MilliQ H_2O in a beaker. The pH of the media was then adjusted to 6.7 by adding NH_4OH. The media was then sterilized by running it through a 0.22 μm vacuum filter. A selected colony from an R1 plate was then added to the media and allowed to incubate in a shake flask for 12 hours. After incubation, OD was recorded to ensure healthy growth, and 10 mL of the culture would be added to the bioreactor for inoculation.

Feed solution was prepared following an AgroSpheres Standard Operating Procedure (SOP). Glycerol and MilliQ H_2O were combined in a beaker and autoclaved. After the solution was allowed to cool, 2M $MgSO_4$, 100X Trace Metal Solution, 5 mg/mL Thiamine, 50 mg/mL Kanamycin were added and the solution was aliquoted into feed bottles.

Bioreactor Set Up

All fermentation studies were performed using Eppendorf BioFlo 120 1L bioreactors. These bioreactors contain sensors which monitor Dissolved Oxygen (DO), pH, and temperature. The reactors also include control mechanisms such as stir rods, gas spargers, hot and cold jackets, and 3 pumps which supply base, antifoam, and substrate. Setting up bioreactors required media

preparation, sterilization, and instrument calibration. Each growth run started with assembling the vessel of the bioreactor and loading in MilliQ H_2O, 10X Citric Phosphate Buffer, and Antifoam. The bioreactors were then autoclaved on the liquid cycle. After autoclaving, the bioreactors were left to cool down to 37 ℃ before being connected to the control panel. The rest of the batch media components were added under a flame which includes 2M $MgSO_4$, 100X Trace Metal Solution, 5 mg/mL Thiamine, 50 mg/mL Kanamycin.

With the batch media prepared and the vessel fully connected to the control panel, the DO and pH sensors must be calibrated and restandardized. The DO probe is unplugged and that measurement is zeroed before it is plugged back in. A sample of the media is taken to measure the pH with an external pH probe, and this measurement is used to restandardize the bioreactor's pH probe. The pumps for the feed solution, NH_4OH, and antifoam are then primed and NH_4OH is added to the media until it reaches a pH of 6.7. The setpoint for the feed pump is then set to zero to initialize it for the feed control algorithm. The feed control algorithm was designed and implemented using NewBrunswick Biocommand Software. After inoculation, a batch is started in the Biocommand software. Cell growth was monitored by taking hourly samples and measuring OD on the BioTek plate reader. After 33-35 hours of running, the reactors are taken down and the culture is stored in a sterile 1 L bottle at 4℃.

RNA Quantification

RNA quantification was performed following an AgroSpheres protocol for a Ribogreen Assay. Before running the assay, all surfaces and instruments are sprayed down with RNase Zap to prevent RNA degradation during the process. Culture cell density was measured using a BioTek plate reader. This measurement is then used to calculate the analytical volume for the assay, which is the volume of culture containing 1e9 cells. This volume was then allocated into 1 mL tubes, with each culture sampled in triplicate. The cells were then digested by adding Trizol to each tube and mixing thoroughly. RNA was then separated from other cell parts by adding chloroform, centrifuging, and extracting the supernatant. Next, the RNA solution was put through two centrifugation wash steps using isopropanol and ethanol. After washing, the RNA pellets were resuspended in RNase-free water, loading dye was added, and each sample was loaded into an SDS-Page gel for gel electrophoresis. The band at 150 bp was then extracted and dissolved on a heat block with a gel dissolving buffer.

To measure fluorescence, each assay sample was loaded into a 96 well plate in triplicate. Prior to loading samples, Ribogreen dye was added to each well at a 40X dilution. A standard curve for fluorescence was also created in the bottom 2 rows of the plate by loading a 10X Ribogreen dilution in the first column and serially diluting by 10X into the next 5 columns. With the samples loaded, fluorescence was measured using a BioTek plate reader. RNA concentration was then calculated through a novel data analysis pipeline which utilized Gen5 and R software. All RNA results were then plotted using the ggplot package on R.

Transformation

The R1 cell line is transformed using electroporation with the appropriate plasmid that allows for the production of both dsRNA and the GFP-fused dsRNA binding protein. Once transformed, the cells are plated on kanamycin-containing agar plates and incubated overnight to grow colonies. In the inducer concentration experiments, colonies were picked directly from the transformation plate to inoculate seed cultures. However, in the induction timing experiments, a single colony from the transformation plate was picked and streaked across another kanamycin-containing agar plate. The new plate streaked with a single colony was then incubated overnight to allow the bacteria to multiply and form more of the exact same colonies. The plate streaked with the single colony was then used to reduce variabilities between colonies on the transformation plate.

Protein Expression

Protein expression began by picking a bacterial colony transformed with the appropriate plasmid to inoculate a 10 mL seed culture. After growing overnight, new expression cultures were inoculated the next morning. All 10 mL cultures contained 10 µL of kanamycin and were inoculated and induced under a sterile flame. The media used in all cultures was also filtered directly into the culture tube using a sterile syringe and filters. Both Arabinose and IPTG were used to induce each culture, and all cultures grew for a period of 5 hours following induction in a 37℃ incubator. After 5 hours, cultures are removed from the incubator and refrigerated overnight before analysis the next day. The protein analysis protocol used can accommodate up to 4 cultures, therefore every experiment run involved varying experimental protein expression conditions in 4 cultures.

Protein Multi-plate Analysis

We created a novel analysis method that uses a BioTek plate reader and Gen5 software program to quantify various metrics related to protein production. Once protein expression was complete, 6 1mL samples were aliquoted from each of the 4 expression cultures into separate microcentrifuge tubes. This resulted in a total of 24 samples, with 6 replicates from each of the 4 cultures to ensure accuracy. The 24 samples were loaded into a 96 well plate in triplicate to measure both the OD_{600} and initial GFP fluorescence. After plate 1 was run, cells were centrifuged for 10 minutes at 14,000 RCF to separate cells from the growth media. The growth media was removed from the samples and cells were resuspended in 200µL of a lysis buffer made up of Bug Buster protein extraction reagent, rLysozyme, benzonase nuclease, deionized water, and Tris-HCl, pH 8.0. The lysis buffer was made fresh for each experiment to ensure efficient lysis and protein extraction, and cells incubated in the lysis buffer for 15 minutes. This was followed by centrifugation for 20 minutes at 16,000 RCF, which separated the soluble fraction from the insoluble fraction. The soluble fraction was the supernatant and was pipetted into fresh microcentrifuge tubes. Before filling the new tubes with the soluble protein fraction, the empty tubes were weighed on an analytical balance to obtain the mass. The filled tubes containing the soluble protein fraction were weighed again to determine the volume of the supernatant in each sample, which was important for future calculations.

Next, a second 96-well plate was loaded with the soluble fractions of all 24 samples in triplicate. Each sample was diluted by a factor of 20 using PBS. In addition, a standard curve was created in the bottom two rows of this plate to ensure accuracy of the GFP fluorescence of the samples. In order to build the standard curve, a stock of purified GFP-fused dsRNA binding protein was serially diluted 6 times. The first dilution was by a factor of 20, and the dilution factor increased by 10 each time. Once loaded, the GFP fluorescence was measured to determine the soluble product protein concentration. The total soluble protein mass was obtained by measuring the total soluble protein concentration of each sample using the Take3 Micro-Volume Plate with the BioTek plate reader. The lysis buffer is used as a blank, and the total soluble protein concentration was determined by the A_{280} measurement.

Protein Data Processing and Analysis

All data from both the first and second plates was consolidated and exported from the Gen5 software into an Excel spreadsheet. The volumes of the soluble protein fraction of each sample were entered into the Excel export and multiplied by the soluble product concentrations obtained from the plate to yield the soluble product protein mass in milligrams. This total soluble protein concentration was multiplied by the volume of the soluble fraction of each sample to find the total soluble protein mass in milligrams. It was important to confirm that total soluble protein expression levels were relatively constant across all cultures during an experiment. Both the soluble product protein mass and total soluble protein mass data was normalized to the OD_{600} of each culture. In all statistical tests, results were considered statistically significant if $p \leq .05$.

Results

Exponential Feed Algorithm

Prior to our project, AgroSpheres' fermentation SOP for the R1 cell line was a manual process which would feed at a constant rate. The feed pump would be turned on by an employee once the cells had exhausted the starting batch media nutrients. When there is a low amount of carbon-source nutrients in the media, cells stop utilizing oxygen for respiration causing the DO of the culture to rise. Therefore, a sudden DO spike at some time between 12-14 hours into the run was an indicator to turn on the feed pump. Once the feed pump was turned on, an employee monitored the culture growth by measuring the OD of the culture every hour. Depending on how well the cells grew, the feed rate would be increased 2-3 times throughout the run.

With these observations in mind, we were able to start designing an automatic feeding algorithm using the Newbrunswick Biocommand Software. The first design requirement of this algorithm was to fully automate the feeding process to make it less labor intensive for AgroSpheres employees. To do this, we needed the algorithm to sense the DO spike. Since the DO of the culture is high at the start of the run (due to a small population of cells), the algorithm needed to wait for at least 12 hours into the run to start monitoring for a spike in the DO trend.

The second design requirement for the algorithm was to feed at a rate which would produce the most biopesticide product. Since minicell production is dependent on parent cell growth, a feed profile which produces a culture with a high cell density and a high RNA concentration would be the optimal feeding strategy. Assuming that the carbon source was the limiting metabolite, it was hypothesized that an exponential feed

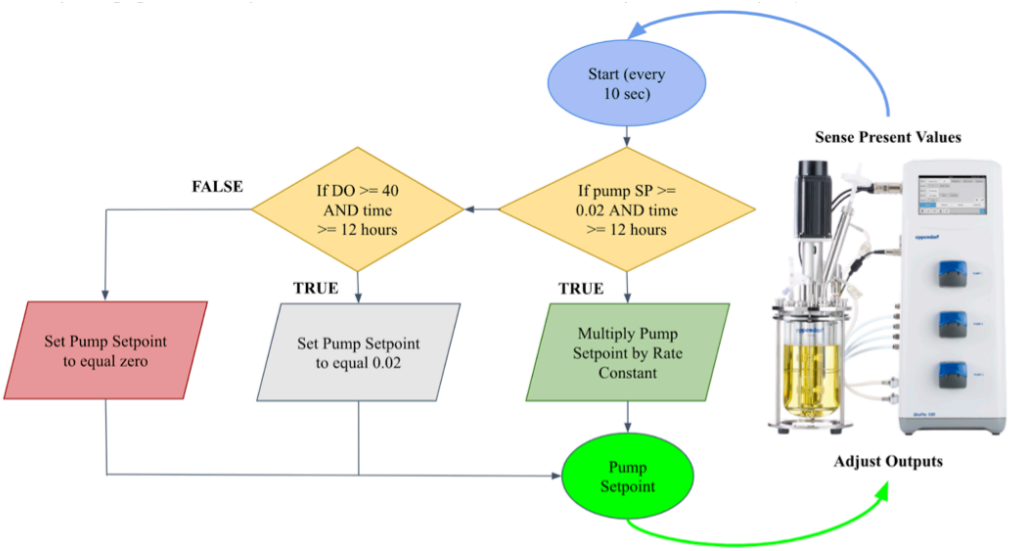


Fig. 1. Diagram of the Exponential Feed Rate Algorithm
The diagram shows the logic structure which is executed every 10 seconds as part of the for loop built into the Biocommannd software. Due to the loop structure, the elif statement (left yellow diamond) executes first to start the feed. This statement waits until 12 hours has elapsed in the run then turns the feed pump on after the DO goes above 40. The first if statement (right yellow diamond) only executed when the pump has been set to .02 mL/hour by the elif statement. Once turned on, the first if statement catches the new setpoint for the feed pump and multiplies this setpoint by an exponential rate constant every 10 seconds. Structuring the algorithm to have the pump's switch in the elif statement ensures that the feed pump will contine to increase exponentially once the DO drops below 40 after the cells start growing again.

profile would meet these criteria since it would provide nutrients proportional to the exponentially growing cell population. Figure 1 shows the final iteration of the algorithm's design cycle.

The next step for implementing this new feeding strategy was to test and validate this algorithm as part of a new fermentation SOP. The first trial run failed when the feed pump did not turn on when the DO spike was observed. A root cause analysis for this run identified a missing step in the SOP which was to initialize the feed pump to 0 before starting the run. To confirm this missing step as the failpoint, we set up the bioreactor and ran a simulation of the run which showed the logic gates were now functioning as intended with the adjustment in place. With this change made, the batch was run again and the feed turned on correctly.

With the feed rate algorithm functioning properly, the next step was to find the optimum rate constant which would maximize the culture's cell density and RNA concentration while maintaining cell viability. We first tested an exponential rate constant of 0.05 ml/hr out of caution to not feed so fast that the bottles would be emptied before we came in the next day to add more. This feed rate was considered to be too slow since it fed nearly half as much glycerol over the time compared to the constant feed rate. As a result, the exponential condition only grew to an OD of 42.6 while the constant condition grew to 75.5. Yet this initial run still showed promising results from the analysis of Product RNA (pRNA) rate of production and yield which can be seen in Figure 2. What we found was that the rate of pRNA production increased by about 58% when fed at an exponential rate of 0.05 mL/hour. Despite the slow growth in the exponential condition, it's higher rate of production made the exponential batch total pRNA yield comparable to the constant rate which is seen in Figure 2B. Additionally, the exponential growth condition fed significantly less base, indicating that the cells may have grown healthier having produced less acetic acid from overflow metabolism.

With these results, the next step was to evaluate higher feed rates for the R1 strain and see if the trend toward efficiency would continue with more cell growth. The rates picked for the next runs were 0.07 and 0.1 mL/hour. The 0.07 mL/hour rate fed a total of 562 mL over the course of 35 hours and reached a max OD of 82.8, a rate of production of 7,379 ng/1e9 cells, and a total pRNA yield of 0.470 µg/mL. The 0.1 mL/hour rate fed a total of 839.5 mL over the course of 35 hours and reached a max OD of 65.7, a rate of production of 7,320 ng/1e9 cells, and a total pRNA yield of 0.370 µg/mL. The results appeared to show that the 0.07 mL/hour rate was the best feed rate tested to that point. However, running the pRNA on a gel showed that the RNA produced in this run was larger than the expected size for pRNA, indicating that a mutated colony was selected for this growth run (see Supplemental Figure 2). As a result, while these optimization runs indicated that the new feeding method was capable of producing a higher yield of pRNA, we could not yet confirm it to be the optimal exponential rate due to the confounding effect of the mutation.

The final step for optimizing the new feed profile was to see if the production obtained at the optimal rate could be replicated and confirmed. To test for confirmation, we performed another growth run with the feed rate set to 0.07 mL/hour. The culture grew to an OD of 41.1, had a rate of pRNA production of 9,284 ng/1e9 cells/OD, and a total pRNA yield of 0.294 µg/mL. Of the non-mutated runs, this run was the least efficient at producing pRNA as seen in Figure 2A. Consequently, this run also had the lowest total pRNA yield of the non-

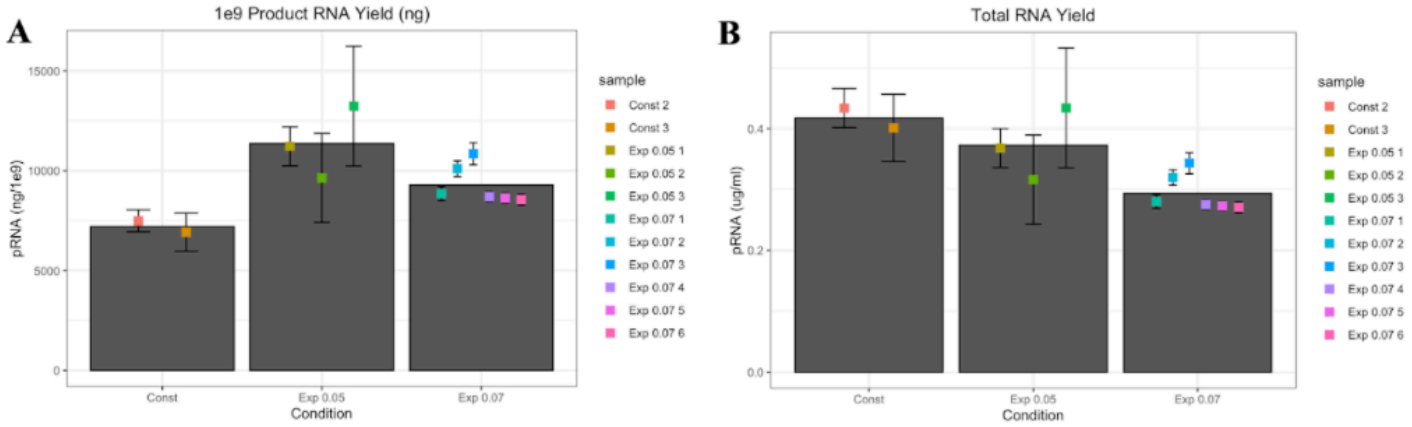


Fig. 2. Bioreactor RNA Quantification. Results of the Ribogreen assay are graphed above with the color representing assay replicates, error bars showing variance in plate reader replicates, and gray bars showing the average of the assay replicates. A) pRNA production rate of the bioreactor culture normalized to 1e9 cells. B) Total concentrations of pRNA produced in the bioreactor culture.

mutated runs as seen in Figure 2B. Despite the low yield, multisizer analysis of the culture showed particle distributions similar to that of cells grown with a constant feed rate (see Supplemental Figure 3). The similarity in the profiles indicated that the new feeding strategy maintained cell viability. The result of this final run indicated that there are still some necessary improvements to be made to this new feeding strategy which will be further examined in the discussion section.

Inducer Concentrations Testing

The first expression variable to be tested was inducer concentration. Agrospheres uses a dual induction system of both Arabinose and IPTG to stimulate overproduction of both the GFP-fused dsRNA binding protein and the target dsRNA sequence that will be used in the RNAi pathway to kill pests. The standard inducer concentrations previously established at AgroSpheres were 10 mM Arabinose and .1 mM IPTG. However, the effects of various inducer concentrations on this particular protein had never been previously characterized by AgroSpheres.

First, expression patterns of the product protein were characterized in response to various inducer concentrations in two different media types. The samples from the first experiment were all grown in Lennox Broth (LB) and induced using a range of concentrations, as noted in Figure 3. The mean soluble product protein masses were 0.0019 mg, 0.0020 mg, 0.0021 mg, and 0.0033 mg respectively for cultures 1, 2, 3 and 4. The highest mean soluble product protein mass was in Culture 4, which was induced using a tenfold increase over the standard inducer concentrations. A one-way ANOVA was performed, followed by Tukey's Range Test post hoc to determine significance between different pairs of groups. The results of the Tukey's Range Testing (see Supplemental Table 1) showed that in culture 4 both the mean soluble product protein mass was statistically significant from all other cultures. Samples for the next experiment were grown in Terrific Broth (TB) and induced using a range of concentrations between the standard concentration and a tenfold increase, as shown in Figure 4. The mean soluble product protein masses were 0.0079 mg, 0.0075 mg, 0.0078 mg, and 0.0080 mg respectively for cultures 1, 2, 3 and 4. A one-way ANOVA was conducted and followed by Tukey's Range Testing. No significant differences in the mean soluble product protein masses were found between any of the cultures (see Supplemental Table 2). The OD_{600} was much higher in all the cultures grown in TB media compared to LB. The mean OD_{600} between all four LB cultures was 1.98, while the mean OD_{600} between all TB cultures was 2.97.

After noticing this difference in soluble product protein expression in the different media types, we set up

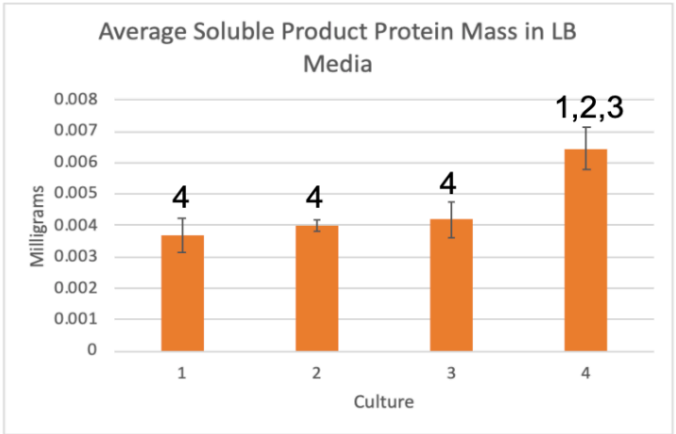


Fig. 3. Mean Soluble Product Protein Mass in LB. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Each culture was grown in LB media and induced using the following concentrations. Culture 1: 0.1 mM Arabinose, 0.001 mM IPTG. Culture 2: 1 mM Arabinose, 0.01 mM Arabinose. Culture 3: 10 mM Arabinose, 0.1 mM IPTG. Culture 4: 100 mM Arabinose, 1 mM IPTG.

several experiments to directly compare the effects of inducer concentrations in LB versus TB. We grew two cultures in TB media, and two cultures in LB media. One culture grown in each different type of media was induced using the standard concentrations, and the other was induced using the tenfold increase over the standard inducer concentrations. The mean soluble product protein masses are shown in Figure 5. Results between cultures grown from different initial bacterial colonies are not comparable due to variabilities in the protein production capabilities of colonies after transformation. The mean soluble product protein masses in cultures 1 and 2 were 0.0035 mg and 0.0032, respectively. An unpaired two tailed t-test instead was used to compare the cultures grown in TB, to each other. The results of this (see Supplemental Table 3) showed that there was no significant difference in mean soluble product protein mass between cultures 1 and 2. The mean soluble product protein masses in cultures 3 and 4 were 0.0054 mg and 0.0061 mg, respectively. Another unpaired two tailed t-test was conducted to compare cultures 3 and 4, grown in LB, to each other. The results (see Supplemental Table 3) indicate that the mean soluble product protein mass was statistically significant in culture 4. Culture 4 had a higher mean soluble product protein mass than culture 3, as shown by the bar graph in Figure 5. In addition, the mean OD_{600} was higher in both the cultures grown in TB compared to cultures grown in LB. The mean OD_{600} between both LB cultures was 2.22, and in the TB cultures was 2.69.

The final inducer concentration experiment was conducted to determine if altering the seed train process created any differences in biomass or protein production. This experiment was set up to grow one seed culture in LB media, which was then used to inoculate two TB cultures, labeled as cultures 1 and 2 in Figure 6. Culture 1 was induced using the standard inducer concentration, and culture 2 was induced using the tenfold increase in inducer concentrations. The mean soluble product protein masses in cultures 3 and 4 were both 0.0023 mg. An unpaired t-test test was used to compare cultures 1 and 2. The results (see Supplemental Table 4) show that there was no significant difference in the mean soluble product protein masses. Another seed culture was grown in TB media and used to inoculate two different LB cultures, labeled as cultures 3 and 4 in Figure 6. Culture 3 used the standard induction concentration, and culture 4 was induced using the tenfold increase in inducer concentrations. The mean soluble product protein masses in cultures 3 and 4 were 0.0015 mg and 0.0028 mg, respectively. Culture 4 had a much higher mean soluble product protein mass than

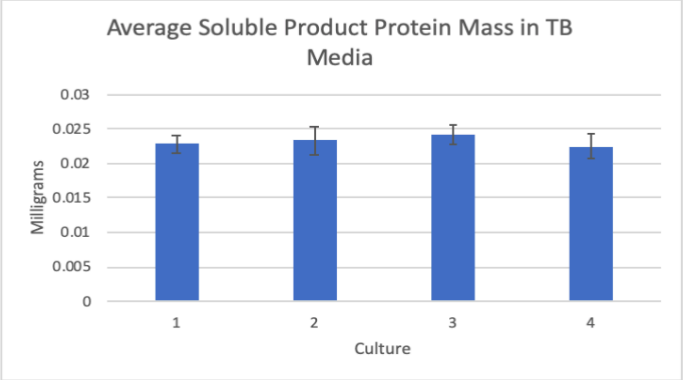


Fig. 4. Mean Soluble Product Protein Mass in TB. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Each culture was grown in TB media and induced using the following concentrations. Culture 1: 10 mM Arabinose, 0.1 mM IPTG. Culture 2: 30 mM Arabinose, 0.3 mM Arabimose **.** Culture 3: 70 mM Arabinose, 0.7 mM IPTG. Culture 4: 100 mM Arabinose, 1 mM IPTG.

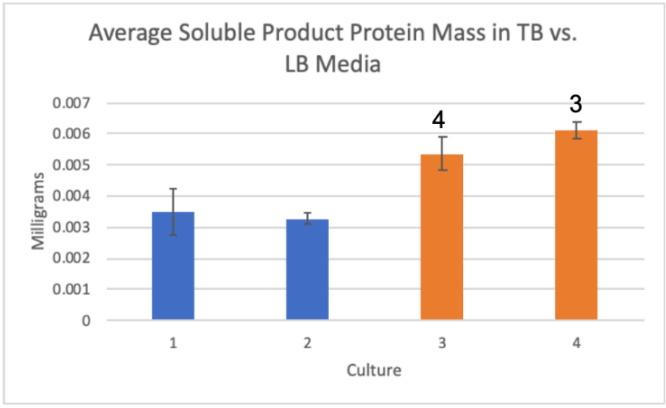


Fig. 5. Mean Soluble Product Protein Mass in TB vs. LB. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Cultures 1 and 2 were grown in TB media, and cultures 3 and 4 were grown in LB media. Cultures 1 and 3 were induced using 10 mM Arabinose, 0.1 mM IPTG. Cultures 2 and 4 were induced using 100 mM Arabinose, 1 mM IPTG.

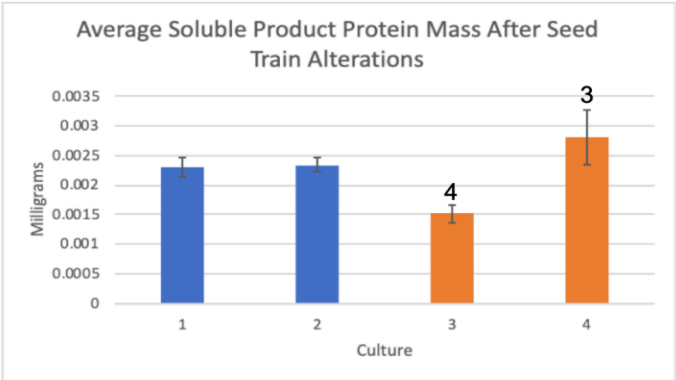


Fig. 6. Mean Soluble Product Protein Mass After Seed Train Alterations. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Cultures 1 and 2 are TB cultures initially inoculated with cells grown in an LB seed culture. Cultures 3 and 4 are LB media cultures initially inoculated with cells grown in a TB seed culture. Cultures 1 and 3 were induced using 10 mM Arabinose, 0.1 mM IPTG. Cultures 2 and 4 were induced using 100 mM Arabinose, 1 mM IPTG.

culture 3. An unpaired two tailed t-test was used to compare cultures 3 and 4 (see Supplemental Table 4), and showed that the mean soluble product protein masses between cultures 3 and 4 were statistically significant. The OD_{600} was higher in cultures 1 and 2, inoculated in TB media compared to cultures 3 and 4, inoculated in LB. The mean OD in cultures 1 and 2 was 3.66, and the mean in cultures 3 and 4 was 2.93.

Induction Timing Testing

The final expression variable that was tested was induction timing. The previously established standard procedure at AgroSpheres was to induce cultures 3 hours after inoculation. After induction, cultures continue to grow for 5 hours and are then removed from the incubator. The timing between inoculation and induction was the time period that was varied, and every culture grew for the standard 5 hours after induction.

The first induction timing experiment was set up to test a wide range of times. Each of the four cultures were grown in TB media and induced using the standard inducer concentrations. Figure 7A shows the mean OD_{600}, which were 1.91, 2.09, 2.43, and 2.55 respectively in cultures 1, 2, 3 and 4. A one way ANOVA followed by Tukey's Range Testing were used to evaluate the cultures (see Supplemental Table 5). Results revealed that the mean OD_{600} of cultures 1 and 2 were significantly different. The mean OD_{600} of cultures 3 and 4 were not significant from each other. However, the OD_{600} in cultures 3 and 4 were higher and statistically significant compared to both cultures 1 and 2. Figure 7B shows the mean soluble product protein masses, which were 0.0037 mg, 0.0033 mg, 0.0030 mg, and 0.0015 mg respectively for cultures 1, 2, 3 and 4. The product protein mass was highest in culture 1 and progressively decreased in cultures 2, 3 and 4. A oone-wayANOVA was conducted and followed by Tukey's Range Testing. The results (see Supplemental Table 6) show no statistically significant differences in mean soluble product protein mass between cultures 1 and 2. There are also no statistically significant differences between the mean soluble product protein masses cultures 2 and 3. However, the mean soluble product protein mass

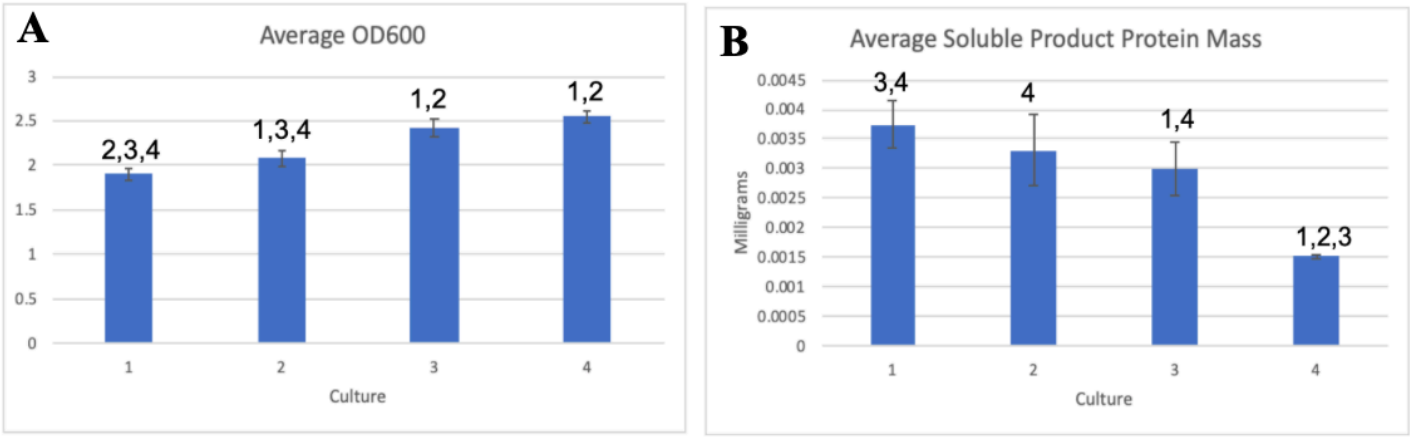


Fig. 7. Mean OD and Soluble Product Protein Mass After Varying Induction Timing. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Culture 1: induced 2 hours after inoculation. Culture 2: induced 3 hours after inoculation. Culture 3: induced 4 hours after inoculation. Culture 4: induced 6 hours after inoculation. A) Bar graph displaying the mean OD. B) Bar graph displaying the mean soluble product protein mass.

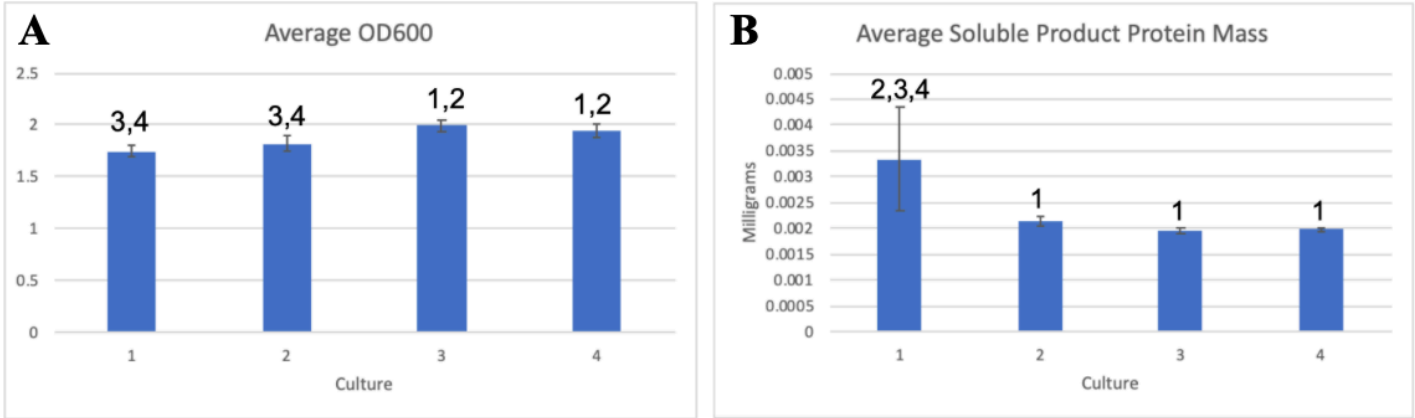


Fig. 8. Mean OD and Soluble Product Protein Mass After Short Range Induction Timing. Error bars represent ± the standard deviation. The numbers above each bar indicate statistical significance from the corresponding numbered culture label on the x-axis. Culture 1: induced 30 minutes after inoculation. Culture 2: induced 1 hour after inoculation. Culture 3: induced 2 hours after inoculation. Culture 4: induced 3 hours after inoculation. A) Bar graph displaying the mean OD. B) Bar graph displaying the mean soluble product protein mass.

in culture 1 was statistically significant compared to cultures 3 and 4.

Based on these results, the next experiment aimed to test a range of shorter induction timing periods. Each of the four cultures were grown in TB media and induced using the standard inducer concentrations. Figure 8A shows the mean OD_{600}, which were 1.75, 1.82, 2.00 and 1.94 respectively in cultures 1, 2, 3 and 4. A one way ANOVA followed by Tukey's Range Testing post hoc showed that the mean OD_{600} was not statistically significant between cultures 1 and 2 or cultures 3 and 4 (see Supplemental Table 7). However, the mean OD_{600} in both cultures 3 and 4 were higher and statistically significant compared to cultures 1 and 2. Figure 8B shows the mean soluble product protein masses, which were 0.0033 mg, 0.0021 mg, 0.0019 mg, and 0.0020 mg respectively for cultures 1, 2, 3 and 4. The product protein mass was highest in culture 1, which was induced first, and decreased in subsequent cultures. A one-way ANOVA and Tukey's Range Testing post hoc was conducted (see Supplemental Table 8), the mean soluble product protein mass in culture 1 was statistically significant compared to cultures 2, 3 and 4. There were no significant differences in mean soluble product protein mass between cultures 2, 3 and 4.

Discussion

The results from the fermentation study show some evidence that further optimization of the new feed strategy will result in a higher amount of pRNA yield for each batch. One improvement of the new design was in fully automating the feeding process. By automatically turning on the feed at the right time, this algorithm removes this task from an AgroSpheres employee's workload. The results from the rate optimization studies found that the new exponential feed was not able to produce pRNA as efficiently as the constant feed rate. However, results from the mutated run at the rate of 0.07 mL/hour indicate that there is potential for this feeding method to produce at or slightly better than the standard feeding method. Ultimately, these results show that further optimization of the feed rate algorithm will likely result in more efficient biopesticide production.

Implementing a fermentation feed control algorithm had also never been attempted before at AgroSpheres. While this study produced evidence that this new feeding method could produce more pRNA, further optimization of the algorithm design would make this process more reliable and efficient. A general issue with any fermentation process is the variability in cell growth due to a multitude of factors during the process. The theory behind the exponential feed rate was to provide exponentially growing cells with proportional nutrients. However, the current algorithm does not have a feedback mechanism to directly respond to the cell growth of the run. One potential future improvement would be to purchase OD sensors for their bioreactors. These sensors are able to track cell growth in real time which would allow the algorithm to sense adjust the feed rate to provide the correct amount of nutrients for the amount of cells currently present in the culture. Adding this feedback mechanism to the algorithm would likely make the feeding process produce pRNA more reliably.

Additionally, AgroSpheres could pursue different options for high throughput fermentation screening. One option would be to invest in a small-scale parallel bioreactor system such as the Ambr 15 (*Ambr® 15 Cell Culture - High Throughput Bioreactor*, n.d.). This system is able to perform growth runs up to 48 15 mL cultures at once, which would generate more data for batch media improvements. Since this system would require a significant investment of company funds, a separate option to optimize the fermentation process would be to outsource the work to a Contracted Research Organization (CRO). These CROs have already invested in more sophisticated fermentation equipment, so collaborating with a technology transfer could be a less expensive option to make improvements to the company's fermentation process.

In addition to the limitations of COVID-19, the fermentation study faced limitations from AgroSphere's production schedule. The fermentation study used the company's bioreactors to test the new feeding algorithm. However, AgroSpheres has a small number of bioreactors which were utilized for higher priority company projects throughout the year. As a result, the fermentation study was limited in its ability to gather enough data to make statistically significant conclusions about improvements made to the process.

The data collected across all of the inducer concentration experiments showed clear expression patterns. In TB media, using the various inducer concentrations did not cause any statistically significant changes in mean soluble product protein production. Based on these results, we concluded that the standard inducer concentrations of 10 mM Arabinose and .1 mM of IPTG are ideal for expression in TB media. However, across all experiments in LB media the mean soluble product protein mass increased and was statistically significant when the inducer concentration was increased tenfold over the standard concentrations. Therefore, the inducer

concentrations of 100 mM Arabinose and 10 mM of IPTG are ideal for maximizing expression in LB media. Another important other factor to consider is the biomass produced in each media type. In all experiments, cultures grown in TB media consistently yielded higher average OD_{600} values compared to the LB grown cultures, indicating that TB is most likely the optimal media type for maximizing cell density.

Characterizing the effects of varying induction timing was another important aspect of improving product protein expression conditions. Cultures induced after 4 hours and 6 hours had a low soluble product protein mass but a high and OD_{600} value compared to the cultures induced after 2 and 3 hours. The culture induced after 30 minutes had a higher mean soluble product protein mass but a much lower OD_{600} in comparison to the cultures induced after 2 and 3 hours. The culture induced after 1 hour had no significant difference in soluble product protein mass, but was lower in OD_{600} compared to the culture induced after 2 and 3 hours. The mean OD_{600} in the first induction timing experiment was higher and statistically significant in the first induction timing experiment between the culture induced after 2 hours and the culture induced after 3 hours. However, this was not repeated in the second experiment because there was no statistically significant difference in OD_{600} measurements between cultures induced 2 and 3 hours after inoculation. In addition, there was no significant difference in mean soluble product protein mass between cultures induced after 2 or 3 hours in either experiment. Therefore, in order to maximize both biomass and soluble product protein production, it is ideal for induction to take place 2 hours after inoculation. The previous standard induction timing of 3 hours after inoculation is not necessary in order to achieve the optimal cell growth and soluble product protein expression.

The expression conditions for the GFP-fused dsRNA binding protein studied throughout the capstone project had never been previously examined at AgroSpheres. However, more research is still needed to further optimize the expression conditions of the product protein. One important consideration is the scalability of this process, because AgroSpheres' plans to expand their production capabilities in the future. We conducted all experiments for this project on a small scale, but testing the expression conditions in a bioreactor will be important for future work. One assumption we made throughout the project was that the improved dsRNA binding protein production was associated with greater dsRNA preservation within the minicells and will result in a greater dsRNA yield. However, future work is still needed to confirm this assumption, and a stability test that compares minicells containing just dsRNA and minicells encapsulating both dsRNA and the GFP-fused dsRNA binding protein. The speed of degradation should be studied to ensure that the product protein does function to extend the longevity of the encapsulated dsRNA. Another future study could involve performing a dual dsRNA and product protein extraction to determine if improvements to the protein expression process result in increased dsRNA yields.

Limitations of this project included time constraints with working within shortened semesters due to COVID-19. In addition, the protein analysis protocol developed was limited in its capabilities. Our original intent was to be able to obtain product protein concentrations in both the soluble and insoluble fractions. However, the insoluble protein concentration proved to be too difficult to read using our plate reader because of the very low GFP fluorescence signal from the protein in the insoluble fraction. In addition, another limitation in our protein experiments was variability between different transformed colonies picked to inoculate cultures. Some colonies performed better in producing protein despite using the exact same expression conditions between experiments. Therefore, data from different experiments could not be compared directly to each other. Only cultures originating from the same initial bacterial colony could be compared directly. This limited our ability to make certain conclusions about which media type and inducer concentration outperformed the other. Despite the limitations, we were able to make improvements to AgroSpheres' platform technology and contribute to their mission to commercialize RNAi biopesticides.

End Matter

Author Contributions and Notes

A.B., J.F. and E.B. designed research, A.B. and E.B. performed research, A.B. and E.B. analyzed data; and A.B. and E.B. wrote the paper.
The authors declare no conflict of interest.

Acknowledgments

We would like to thank the entire AgroSpheres team for their support in this project. We would like to especially thank our advisor Joseph Frank, Liz Somers and Jacob Englaender.

References

Alavanja, M. C. R. (2009). Pesticides Use and Exposure Extensive Worldwide. *Reviews on Environmental Health*, *24*(4), 303–309.

Ambr® 15 Cell Culture—High Throughput Bioreactor. (n.d.). Sartorius. Retrieved May 6, 2021, from https://www.sartorius.com/en/products/fermentation-bioreactors/ambr-multi-parallel-bioreactors/ambr-15-cell-culture

Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., & Hwa, T. (2015). Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature*, *528*(7580), 99–104. https://doi.org/10.1038/nature15765

Coman, G., Farcas, A., Matei, A. V., & Florian, C. (2013). Pesticides Mechanisms of Action in Living Organisms. In L. I. Simeonov, F. Z. Macaev, & B. G. Simeonova (Eds.), *Environmental Security Assessment and Management of Obsolete Pesticides in Southeast Europe* (pp. 173–184). Springer Netherlands. https://doi.org/10.1007/978-94-007-6461- 3_16

Fletcher, S. J., Reeves, P. T., Hoang, B. T., & Mitter, N. (2020). A Perspective on RNAi-Based Biopesticides. Frontiers in Plant Science, 11. https://doi.org/10.3389/fpls.2020.00051

Islam, M. T., Davis, Z., Chen, L., Englaender, J., Zomorodi, S., Frank, J., Bartlett, K., Somers, E., Carballo, S. M., Kester, M., Shakeel, A., Pourtaheri, P., & Sherif, S. M. (2021). Minicell-based fungal RNAi delivery for sustainable crop protection. *Microbial Biotechnology*, *n/a*(n/a). https://doi.org/10.1111/1751-7915.13699

Kim, D. H., & Rossi, J. J. (2008). RNAi mechanisms and applications. BioTechniques, 44(5), 613–616. https://doi.org/10.2144/000112792

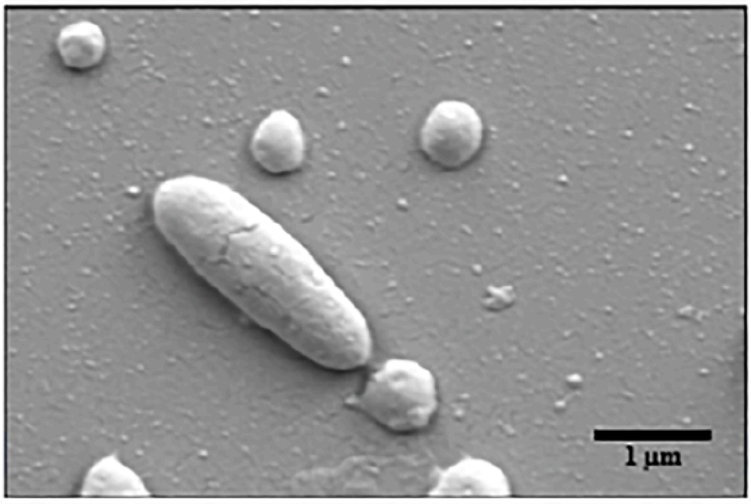
Mamta, B., & Rajam, M. V. (2017). RNAi technology: A new platform for crop pest control. *Physiology and Molecular Biology of Plants*, *23*(3), 487–501. https://doi.org/10.1007/s12298-017-0443-x

Mnif, W., Hassine, A. I. H., Bouaziz, A., Bartegi, A., Thomas, O., & Roig, B. (2011). Effect of Endocrine Disruptor Pesticides: A Review. International Journal of Environmental Research and Public Health, 8(6), 2265–2303. https://doi.org/10.3390/ijerph8062265

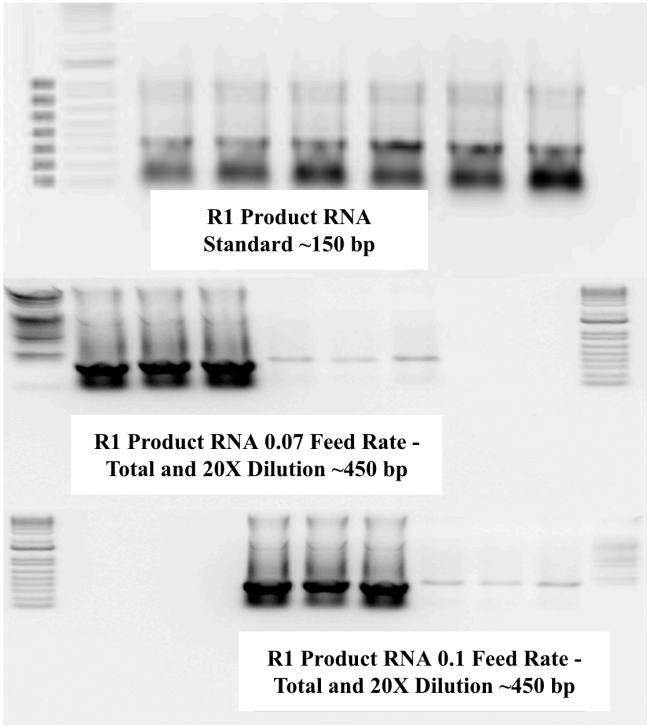
Nicolopoulou-Stamati, P., Maipas, S., Kotampasi, C., Stamatis, P., & Hens, L. (2016). Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture. *Frontiers in Public Health*, *4*. https://doi.org/10.3389/fpubh.2016.00148

Sanborn, M., Kerr, K. J., Sanin, L. H., Cole, D. C., Bassil, K. L., & Vakil, C. (2007). Non-cancer health effects of pesticides. Canadian Family Physician, 53(10), 1712–1720.

Supplemental Material

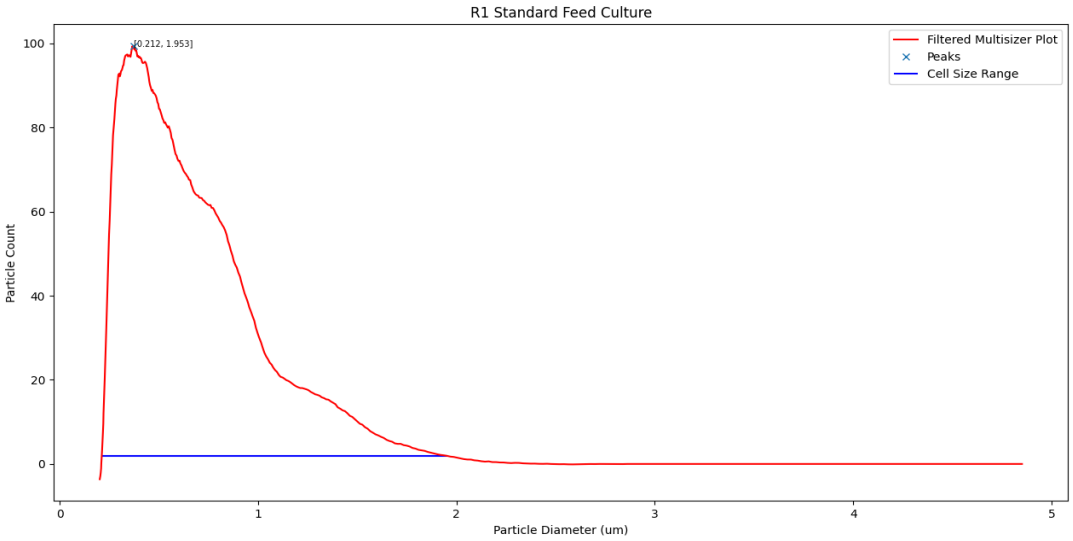


Supplemental Fig. 1. pRNA Gel Electrophoresis. Shows bacterial minicells compared to the parent cell.

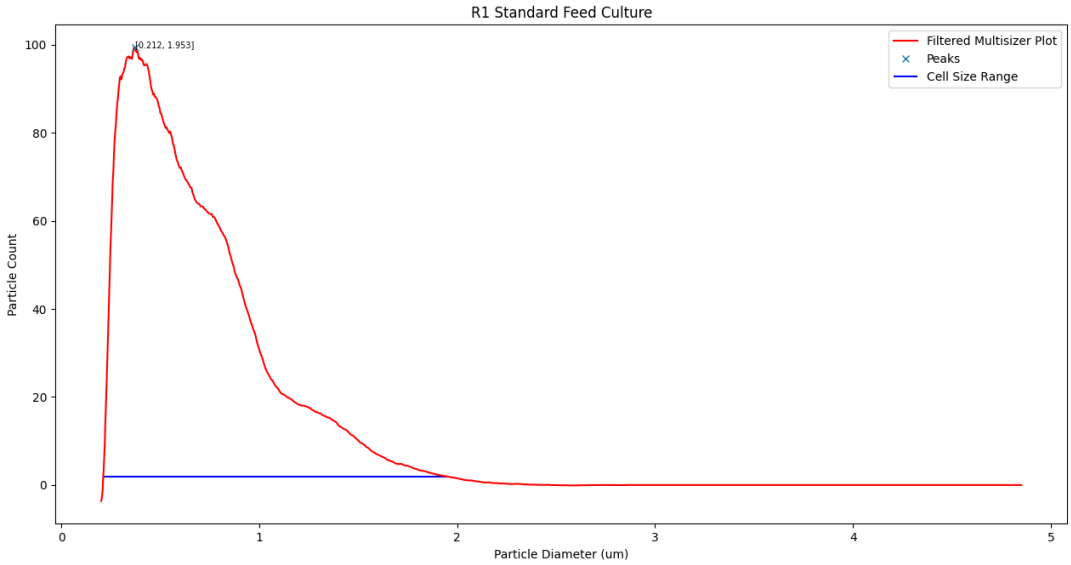


Supplemental Fig. 2. pRNA Gel Electrophoresis. The bottom two images show pRNA from mutated bioreactor cultures which appear to be at ~450 bp. This is larger than the normal pRNA length for the R1 cell line seen in the top image.

A



B



Supplemental Fig. 3. Multisizer Profile Analysis. Multisizer analysis measures the particle size distribution for the bioreactor culture. This allows us to see if there is abnormal cell morphology or cell break down. A) R1 cells line grown at a constant feed rate. This is considered a health size distribution. B) The profile of R1 cells grown at an exponential feed rate of 0.07 mL/hour

Supplementary Table 1. Results of Tukey's Range Testing in LB Media.
Comparisons of the mean soluble product protein mass in cultures grown in LB media.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.8867	No
Culture 1	Culture 3	0.6666	No
Culture 1	Culture 4	3.966E-07	Yes
Culture 2	Culture 3	0.9740	No
Culture 2	Culture 4	1.609E-06	Yes
Culture 3	Culture 4	3.760E-06	Yes

Supplemental Table 2. Results of Tukey's Range Testing in TB Media.
Comparisons of the mean soluble product protein mass in cultures grown in TB media.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.6610	No
Culture 1	Culture 3	0.9896	No
Culture 1	Culture 4	0.9874	No
Culture 2	Culture 3	0.8237	No
Culture 2	Culture 4	0.4747	No
Culture 3	Culture 4	0.9193	No

Supplemental Table 3. Results of Unpaired T-Tests in LB and TB Media.
Comparison of the mean soluble product protein mass in cultures grown in LB and TB media.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.1317	No
Culture 3	Culture 4	0.0002	Yes

Supplemental Table 4. Results of Unpaired T-Tests After Altering the Seed Train. Comparison of the mean soluble product protein mass in cultures grown using the altered seed train events.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.5978	No
Culture 3	Culture 4	0.00007	Yes

Supplemental Table 5. Results of Tukey's Range Testing for OD After Altering Induction Timing. Comparisons of the mean OD in cultures induced between 2 and 6 hours after inoculation.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.0050	Yes
Culture 1	Culture 3	3.103E-09	Yes
Culture 1	Culture 4	7.158E-11	Yes
Culture 2	Culture 3	2.983E-06	Yes
Culture 2	Culture 4	2.290E-08	Yes
Culture 3	Culture 4	0.07094	No

Supplemental Table 6. Results of Tukey's Range Testing for Protein Mass After Altering Induction Timing. Comparisons of the mean soluble product protein mass in cultures induced between 2 and 6 hours after inoculation

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.3413	No
Culture 1	Culture 3	0.0330	Yes
Culture 1	Culture 4	1.111E-07	Yes
Culture 2	Culture 3	0.5845	No
Culture 2	Culture 4	2.959E-06	Yes
Culture 3	Culture 4	4.388E-05	Yes

Supplemental Table 7. Results of Tukey's Range Testing for OD after Short Range Induction Timing. Comparisons of the mean OD in cultures induced between 30 minutes and 3 hours.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.1766	No
Culture 1	Culture 3	2.515E-06	Yes
Culture 1	Culture 4	8.0967E-05	Yes
Culture 2	Culture 3	0.0002	Yes
Culture 2	Culture 4	0.0101	Yes
Culture 3	Culture 4	0.3793	No

Supplemental Table 8. Results of Tukey's Range Testing for Protein Mass after Short Range Induction Timing. Comparisons of the mean soluble product protein in cultures induced between 30 minutes and 3 hours.

Group 1	Group 2	P-Value	Statistically Significant
Culture 1	Culture 2	0.0030	Yes
Culture 1	Culture 3	0.0006	Yes
Culture 1	Culture 4	0.0009	Yes
Culture 2	Culture 3	0.9059	No
Culture 2	Culture 4	0.9468	No
Culture 3	Culture 4	0.9992	No