REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
12-03-2019	Final Report	1-Sep-2017 - 30-Nov-2018

4. TITLE AND SUBTITLE
Final Report: Competence machinery of Streptococcus pneumoniae as a way for in vivo de novo DNA sequencing

5a. CONTRACT NUMBER
W911NF-17-2-0132

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Lausanne
Biophore

00000 -

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
71618-LS-DRP.1

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jan-Willem Veening
UU	UU	UU	UU		19b. TELEPHONE NUMBER 000-000-0000

RPPR Final Report
as of 25-Apr-2019

Agency Code:

Proposal Number: 71618LSDRP **Agreement Number: W911NF-17-2-0132**
INVESTIGATOR(S):

Name: Jan-Willem Veening
Email: Jan-willem.veening@unil.ch
Phone Number: 0000000000
Principal: Y

Organization: **University of Lausanne**
Address: Biophore, Lausanne 1015, 00000
Country: CHE

DUNS Number: 482392628 EIN:
Report Date: 28-Feb-2019 Date Received: 12-Mar-2019
Final Report for Period Beginning 01-Sep-2017 and Ending 30-Nov-2018
Title: Competence machinery of Streptococcus pneumoniae as a way for in vivo de novo DNA sequencing
Begin Performance Period: 01-Sep-2017 **End Performance Period:** 30-Nov-2018
Report Term: 0-Other

Submitted By: Lance Keller Email: lanceedward.keller@unil.ch
Phone: (004) 178-7190682

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: **STEM Participants:**

Major Goals: The goal of this proposal is to design a system capable of DNA sequencing that utilizes in vivo cells. It is theorized that with TIRF microscopy it is possible to read a FRET signal as labeled DNA is imported into the bacterium Streptococcus pneumoniae using native transporter machinery that has been labeled with a corresponding fluorescent protein. The pneumococcal DNA uptake system always proceeds in a 3´ to 5´ fashion making downstream analysis and assembly easier. Along with this advantageous property, roughly 100 kb can be imported by a single pneumococcal DNA transporter in under 20 minutes. While the DNA is fragmented into a ssDNA molecule once entering the cell, the median intracellular fragment size still remains >6 kb. After an eclipse period of roughly 20 minutes, additional DNA can be imported by the same cell. Therefore, 1Mb can be imported by a single cell through one importer in under seven hours. Couple this with thousands of cells in unison, which are only 1-2 microns in length, it would be possible to scale this system to any size desired to allow extremely high through-put sequencing. This method could revolutionize the method of genome sequencing, which is a technique that is used worldwide with increasing frequency. The ability to grow sequencing chips could reduce the overall cost of sequencing and combined with longer read lengths will make downstream analysis easier and increase sequencing availability. Sequencing of eukaryotic genomes would become much easier because of the potential scalability of the proposed system and the read lengths which help overcome assembly of repetitive regions in genomes. This would make the field of personalized medicine much more feasible potentially aiding in general human health.

Phase I of the proposed project, which can be completed within one year of the start would attempt to deliver an in vivo system capable of producing a FRET signal during DNA importation. During Phase I the production of DNA transport machinery that has viable fluorescent tags of various proteins will be produced. Along with this, Phase I will also deliver which DNA labels and dyes are able to produce viable FRET signals with specific fluorescent proteins during importations. The proposed final deliverable for Phase I would be a nonvirulent strain of S. pneumoniae with fluorescent tags on the native DNA importer that have a known correspondence to a specific DNA label that produces a viable FRET signal with the possibility of distinguishing between differences in incoming nucleotides. Phase I would show the viability of producing a commercial product in the future that would use a specialized microfluidics system coupled with a fluorescence output reader, synthetic bacteria or liposomes, and labeled DNA from any source.

Deliverables for this project include
1. Genetically engineered S. pneumoniae (without capsule so Biosafety level 1) strains containing functional fluorescent protein fusions to proteins of the transformation machinery.

2. Strains made in (1) that contain a single DNA 'sequence' channel per cell.
3. An optimized and generic method to fluorescently label DNA from any source
4. Detection of FRET between fluorescently labeled DNA and fluorescently labeled pneumococci

This project is aticipated to last 12 months with the procedding milestones as checkpoints

1) 12 Weeks: Functional fusions of various competence components and strains lacking single deletions at genes of interest. Known level of induction required to produce single functional DNA import channels.

2) 22 Weeks: Strains with multiple fluorescent proteins tagged to various components of the DNA import system of the pneumococcus. Strains that produce these tagged proteins at a single transporter per cell level without induction.

3) 32 Weeks: A usable microfluidics setup capable of real time observation of DNA import into the pneumococcal cell. Knowledge of which labeled transporter machinery is capable of producing FRET signal due to proximity of incoming DNA.

4) 38 Weeks: A combination of known fluorescent proteins and DNA labels that produce the best signal for downstream interpretation.

5) 42 Weeks: With information from previous tasks it will be possible to know which transporters should be labeled with which specific protein and what labels should be used on incoming DNA to produce viable FRET signals that could distinguish between the different incoming nucleotides.

6) 50 Weeks: Information of the FRET signal for importation of specific nucleotides and changes in signal gain as the FRET signal changes between nucleotides.

The deliverables for the project were completed with numerous engineering pneumococcal strains being created. The project was altered to focus on detection of a FRET signal between adjacently labeled proteins instead of a signal between the protein and incoming DNA. While focusing on this aspect we were able to detect a FRET signal between adjacent proteins as well as measuring changes in the FRET signal when DNA is being imported.

Six milestones during this project were achieved and completed, but an additional three months were utilized to complete the project.

Milestone 1 for the first 12 weeks was completed on time in mid November of 2017. Milestone 2 was also completed on time by the end of December 2017. Milestones 3 and 4 were delayed from expected completion time and completed at the end of March and May of 2018 respectively. The final two milestones were also delayed in completion being completed at the end of September and November 2018 respectively. While there were delays in completing the milestones and deliverables we were able to complete this tasks as anticipated with the alterations in the approved plan by using interprotein FRET as opposed to protein-DNA FRET.

Accomplishments: Stated goals of the project:
The project was divided into six tasks with two to three subtask for each task. The entire project was to be done by the primary researcher with no subcontractor. The project was broken down into quarters with an additional no-cost extension of three months added to the end of the project. Task 1 consisted of five subtasks, which were creating a capsule and pneumolysin deficient strain, create C- and N-terminal fluorescent protein fusions to components of the DNA transporter machinery, deletion of targeted DNA transporter genes, test functionality of DNA transporter protein fusions, and determine optimal induction level of promoter for single molecule expression. Task 2 includes creating new fusions with different fluorescent proteins, altering inducible promoter, and determining promoter for single transporter level. A revision was made for altering the inducible promoter in task 2.2. In lieu of this task we decided to create a FRET pair between adjacent proteins of the transporter machinery. This was done in an attempt to determine if a variance in the FRET signal can be recorded that can relate to a specific nucleotide being imported as a method of DNA sequencing. Task 3 consists of setting up a microfluidics system for visualization of DNA importation, fluorescently labeling of DNA, and initial recording of FRET signal. Task 4 was designed to test different fluorescent proteins and various DNA stains for optimizing FRET signal. Task 5 was to visualize multiple foci simultaneously and further optimize FRET combinations. Task 6 was to use data from task 5 to label

synthesized DNA to get base line FRET signal for a specific incoming nucleotide.

Task one created numerous constructs for use within the pneumococcus and an avirulet strain of pneumococcus. Testing of fluorescent proteins fusions to all major pneumococcal DNA transporters was accomplished and viability of signal for both carboxy and amine terminal fusions determined. The functionality of these fusions was also determined after fusion to fluorescent proteins.

Task two expanded on task one by creating even more options for fluorescent protein fusions, greatly expanding our ability to complete the project. This task also produced the first viable FRET signal between adjacent labeled proteins.

Task three focused on optimization of a new sequencing chip developed specifcially for this project as well as determining the best methods of DNA labeling techniques for use within our system.

Task four focused on optimization of fluorescent proteins being used for the project as well as determining the best fluorescent proteins to use for a FRET combination. This resulted in several new genetic constructs and the creation of a novel fusion protein technique.

Task five examined the interactions between labeled proteins and DNA. We found that the signal between adjacently labeled proteins was significantly higher than that of labeled protein interacting with labeled DNA. We were also able to make measurements of FRET signal between labeled proteins when DNA was added to the system.

Task six was the culmination of the project to determine if specific nucleotides being imported impacted the FRET signal in different ways. We see that when compared to baseline signal there is a significant alteration when DNA is being imported and that the signal does vary when different types of DNA are being imported.

Training Opportunities: Nothing to Report

Results Dissemination: A peer-reviewed scientific article is being prepared for submission that outlines the creation of the cloning vectors created through this project.

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI
Participant: Jan-Willem Veening
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Lance Edward Keller
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

RPPR Final Report
as of 25-Apr-2019

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Anne-Stephenie Rueff
Person Months Worked: 14.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Competence machinery of *Streptococcus pneumoniae* as a way for *in vivo de novo* DNA sequencing: Final Report (September 1st 2017 through November 30th 2018)

Brief summary of the project:

The current goal for the first year of this cooperative agreement is to determine the viability of obtaining FRET signal from exogenous DNA upon interaction of the DNA transport machinery of *Streptococcus pneumoniae*. The long term goal would be to determine the possibility of defining the incoming DNA nucleotide sequence based on FRET signal obtained as a means of sequencing DNA.

Stated goals of the project:

The project was divided into six tasks with two to three subtask for each task. The entire project was to be done by the primary researcher with no subcontractor. The project was broken down into quarters with an additional no-cost extension of three months added to the end of the project. Task 1 consisted of five subtasks, which were creating a capsule and pneumolysin deficient strain, create C- and N-terminal fluorescent protein fusions to components of the DNA transporter machinery, deletion of targeted DNA transporter genes, test functionality of DNA transporter protein fusions, and determine optimal induction level of promoter for single molecule expression. Task 2 includes creating new fusions with different fluorescent proteins, altering inducible promoter, and determining promoter for single transporter level. A revision was made for altering the inducible promoter in task 2.2. In lieu of this task we decided to create a FRET pair between adjacent proteins of the transporter machinery. This was done in an attempt to determine if a variance in the FRET signal can be recorded that can relate to a specific nucleotide being imported as a method of DNA sequencing. Task 3 consists of setting up a microfluidics system for visualization of DNA importation, fluorescently labeling of DNA, and initial recording of FRET signal. Task 4 was designed to test different fluorescent proteins and various DNA stains for optimizing FRET signal. Task 5 was to visualize multiple foci simultaneously and further optimize FRET combinations. Task 6 was to use data from task 5 to label synthesized DNA to get base line FRET signal for a specific incoming nucleotide.

Task One:

Completion of subtask 1.1 was accomplished by the allelic replacement of two of the main virulence determinants of *Streptococcus pneumoniae*, the polysaccharide capsule and the hemolytic toxin pneumolysin. Previous research has shown that these two factors are a main component to pneumococcal virulence and deletion of these genes produces a nonvirulent strain that can be safely handled in any environment. Also accomplished was subtask 1.2, which was to create fluorescently fused DNA transporter proteins. Through this task two distinct new cloning vectors were created. Both vectors utilized an IPTG inducible promoter with one vector being designed for rapid C-terminal mNeongreen fusions of targeted genes and the other for N-

terminal fusions. With these vectors, amplification of targeted genes followed by either restriction digestion cloning (C-ter fusions) or through Golden Gate Assembly (N-ter fusions) can be rapidly achieved (Figure 1).

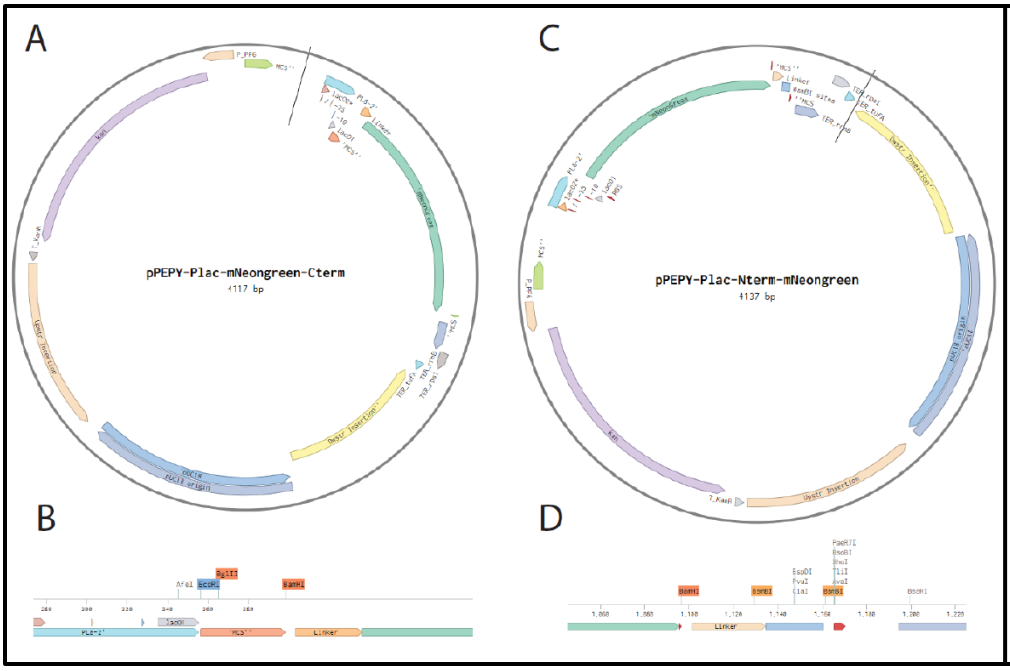

Figure 1: New chromosomal integration plasmids for terminal fusions with fluorescent protein mNeongreen. A) Plasmid map for C-terminal fusion of mNeongreen. B) Restriction digestion method for rapid cloning with C-terminal fusion plasmid. C) Plasmid map for N-terminal fusion of mNeongreen. D) Restriction digestion method for rapid cloning with N-terminal fusion plasmid.

Using these new constructs seven DNA transporter proteins from the pneumococcal competence machinery were chosen (on the basis of literature research) to be fused to mNeongreen, these proteins are ComGA, ComGB, ComGC, ComEC, ComEA, ComFA, and EndA. The group was able to successfully fuse these proteins, excluding EndA, with a C-terminal and N-terminal linkage of mNeongreen (Figure 2).

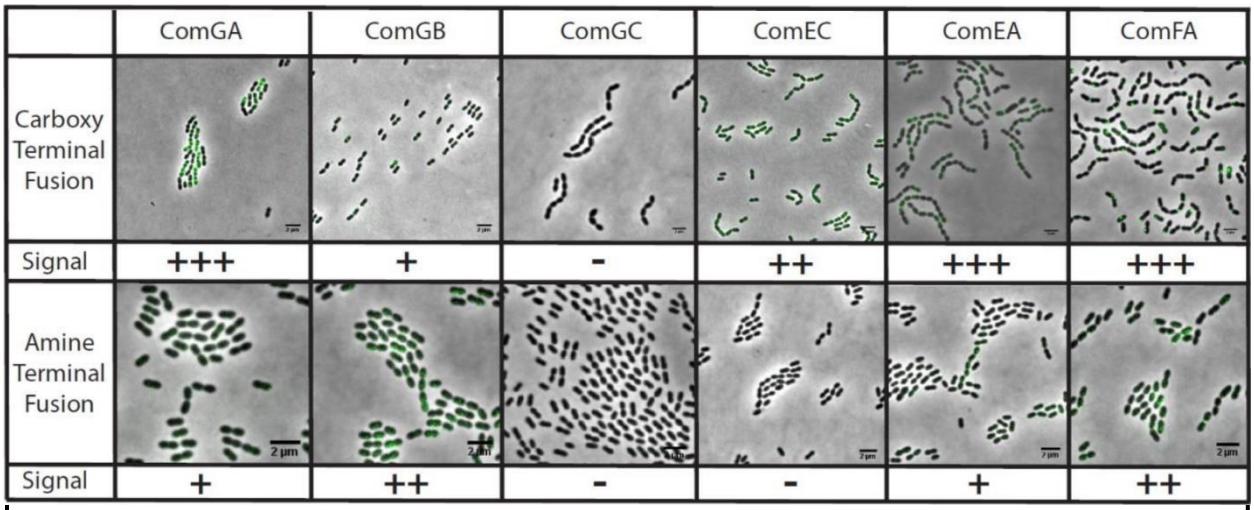

Figure 2: Examples of the six DNA transporter machinery proteins successfully fused at both carboxy (C) terminus and amine (N) terminus with mNeongreen. Signal strength is either weak (-) or strong (+/+++) based on signal intensity and localization of signal.

Subtask 1.3 and 1.4 consists of gene deletions of all transporter components of interest and testing of the functionality of the labeled components. This was done through allelic replacement of the genes of interest with a selectable marker. Testing for functionality was done at various levels of induction, either 0, 5, 100, or 1000 µM of inducer IPTG. All deletions were compared to wildtype strain D39 as well as a strain containing both the mNeongreen fusion and the native protein. As seen in Figure 3 the five gene deletions with their respective induced fusions were transformable with the exception of ComFA. There was a graded response of transformability based on the amount of inducer added. Expression of fusion protein with the native gene intact reduced the transformability of certain strains, notably ComGB, which blocked transformation at high induction level.

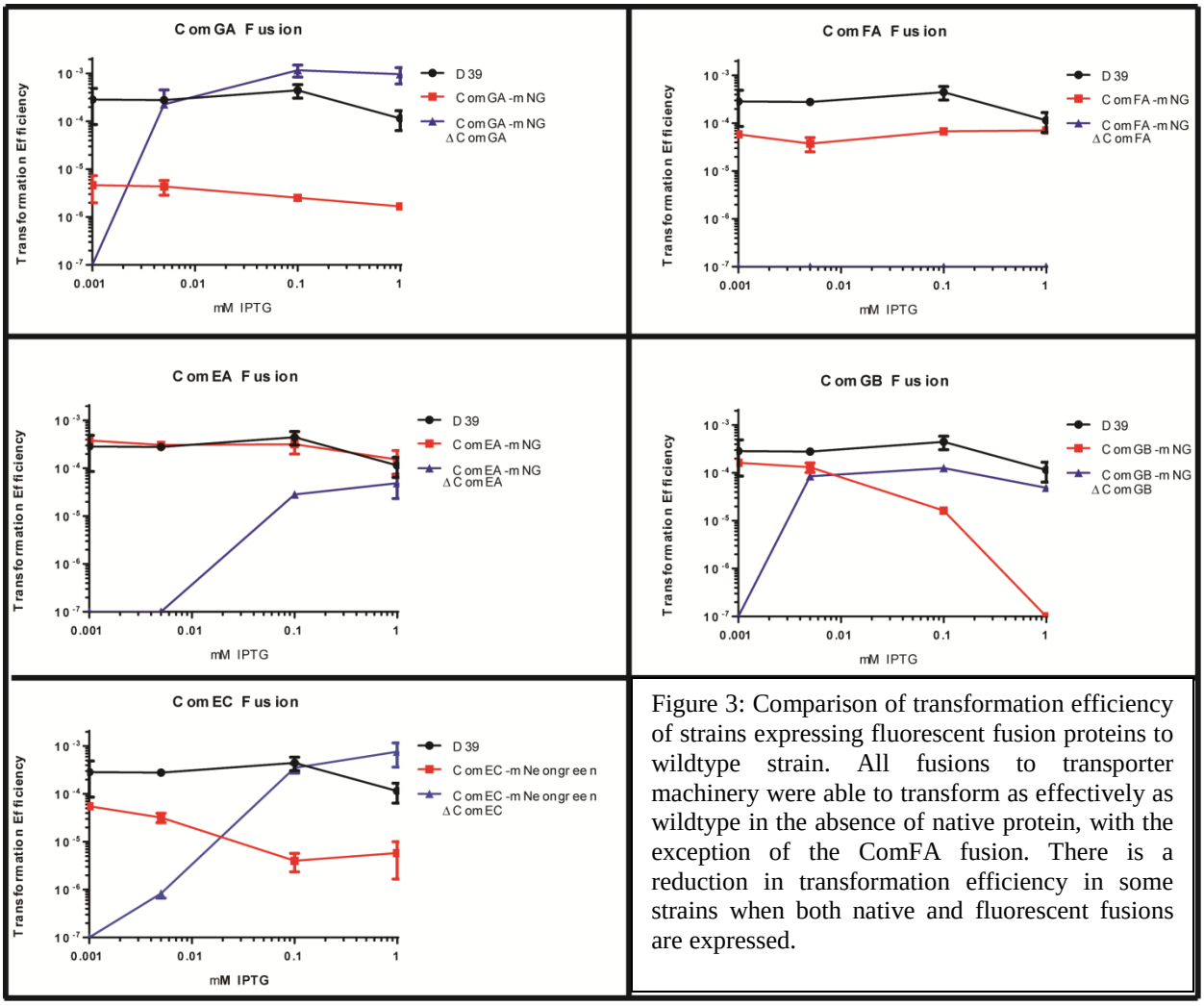

Figure 3: Comparison of transformation efficiency of strains expressing fluorescent fusion proteins to wildtype strain. All fusions to transporter machinery were able to transform as effectively as wildtype in the absence of native protein, with the exception of the ComFA fusion. There is a reduction in transformation efficiency in some strains when both native and fluorescent fusions are expressed.

Based on our observation the three most likely candidates for future study are the ComEA, ComFA, and ComGA proteins based on localization pattern and signal strength. Using our inducible promoter we were able to complete subtask 1.5 to determine optimal induction for expressing fusion proteins to the lowest detectable level (Figure 4).

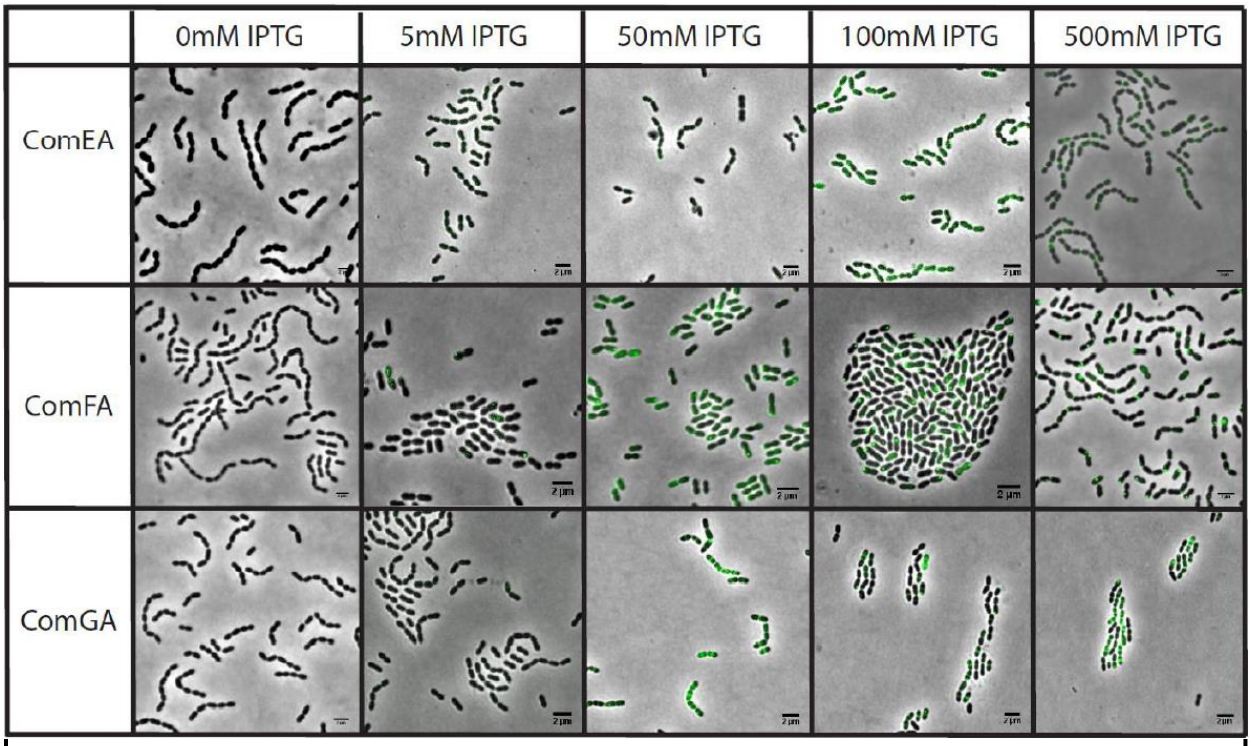

Figure 4: The three most promising proteins are presented here in greater detail and details the range of inducing agents and varying expression that can be achieved.

Task Two:

The main objective for this task was to alter the fluorescent proteins on the fusion vectors and alter the promoters of the vectors to allow for single fluorescent foci. This task was altered due to the presence of one to two foci with the current promoter being used. The second subtask was altered to try and combine multiple labeled fluorescent transporter proteins. This was done to try and create a system to allow for visualization of FRET between adjacent proteins and determine if a variance in the FRET signal can be recorded that can relate to a specific nucleotide being imported as a method of DNA sequencing. For the creation of fusion proteins with various fluorescent labels, four new cloning vectors were created. These cloning vectors are based on the previously constructed vectors for N and C terminal fusion with mNeongreen, but with the fluorescent gene exchanged with either mTurquoise2 or mScarlet-I creating four new cloning vectors. With these vectors, fusions of both ComGA and ComFA were created (Figure 5).

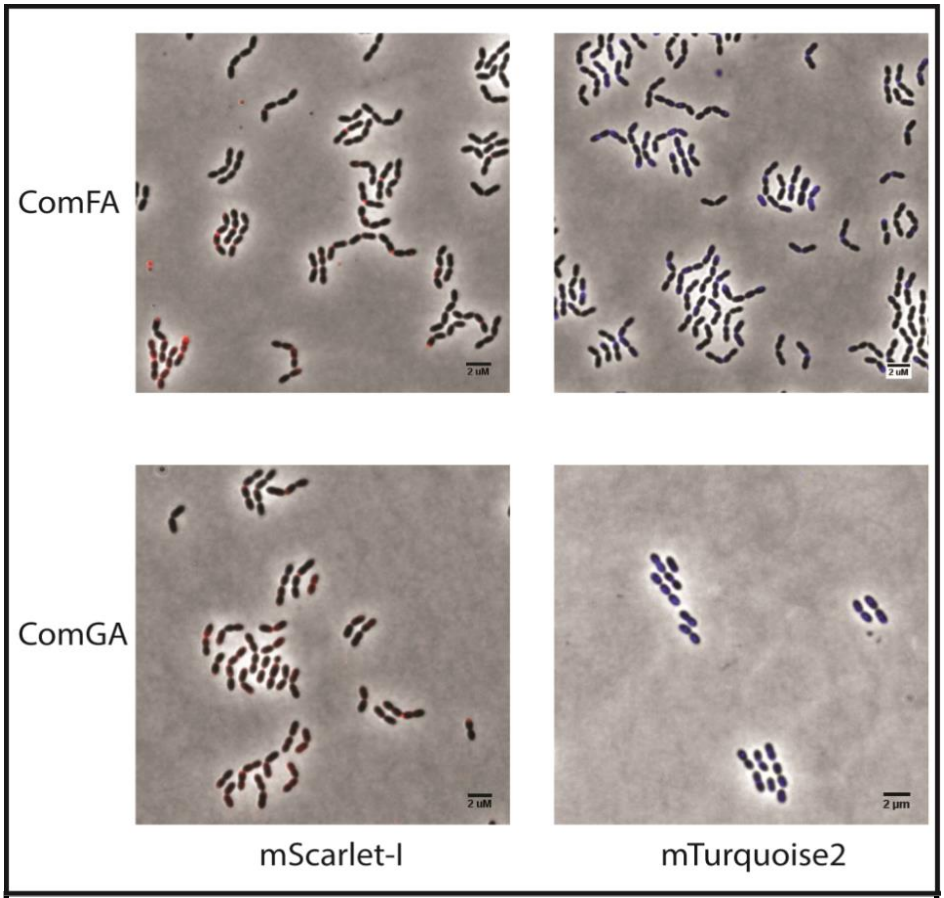

Figure 5: Pneumococcal strains expressing new fluorescent protein fusions in a manner similar to previous mNeongreen fusion. Both ComGA and ComFA were fused to mScarlet-I and mTurquoise2.

As mentioned earlier task 2.2 and 2.3 were revised to attempt a second method of DNA sequencing. To do this a second cloning vector was created that relies on a different integration site into the pneumococcal chromosome, this allows single copy expression of two different genes simultaneously. Upon expression of two labeled proteins that co-localize both distinct signals are obtainable. In addition to the individual fluorescent signal, potential FRET signal is observed. This is obtained by exciting the donor fluorescent protein and reading the emission from the acceptor fluorescent protein. A problem when performing FRET analysis is that often the donor fluorescent protein has overlapping emission with the acceptor protein. To accommodate for this the background fluorescence from the donor must be subtracted from the FRET signal leaving behind fluorescence from the acceptor. When this was done we were able to obtain signal that was over the background from the donor and can be from fluorescence from the acceptor. The red fluorescent protein does not fluoresce in FRET settings (Figure 6, red arrows). The FRET signal obtain is roughly 10% of the fluorescence obtained from excitation of the red fluorescent protein.

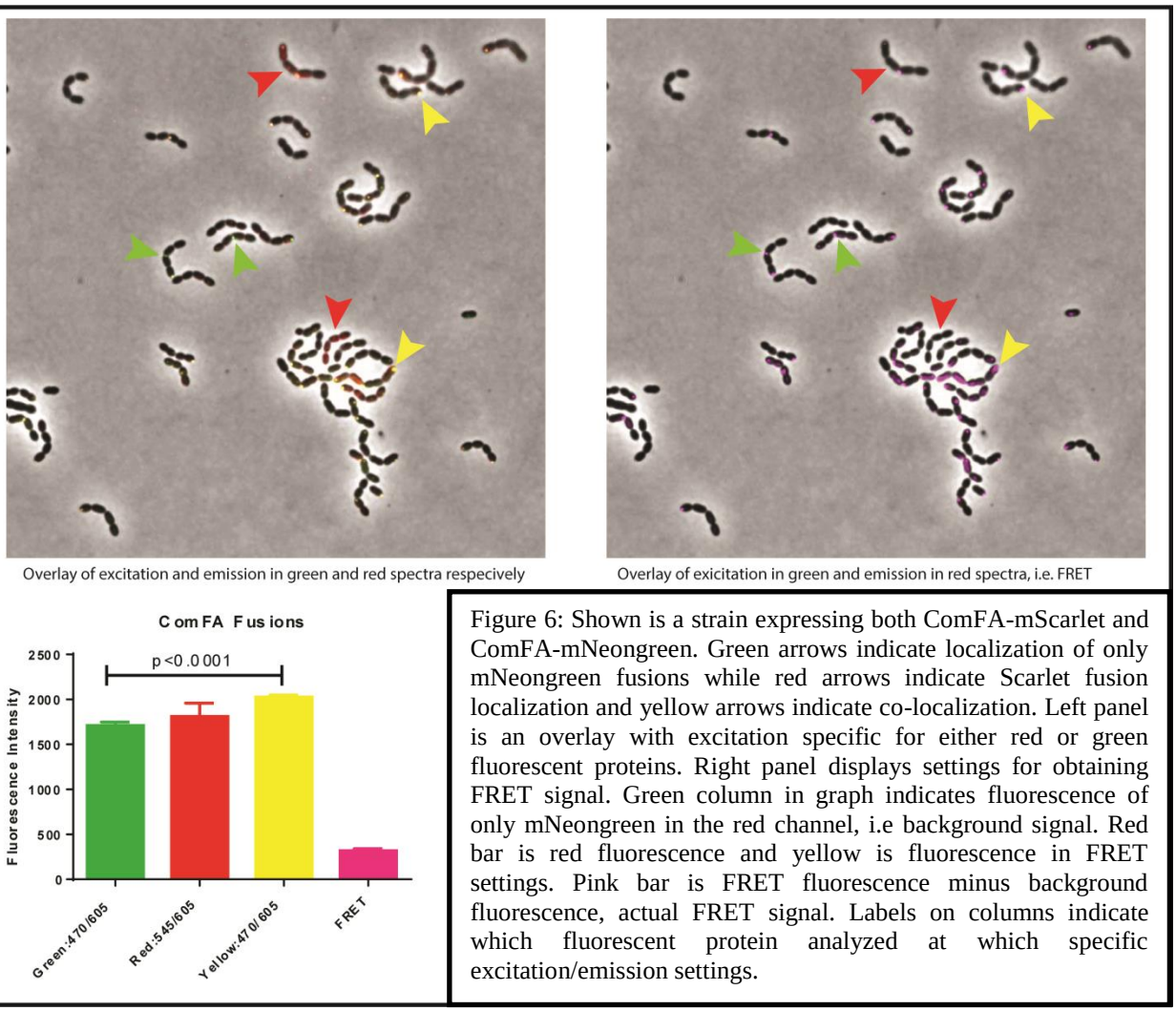

Figure 6: Shown is a strain expressing both ComFA-mScarlet and ComFA-mNeongreen. Green arrows indicate localization of only mNeongreen fusions while red arrows indicate Scarlet fusion localization and yellow arrows indicate co-localization. Left panel is an overlay with excitation specific for either red or green fluorescent proteins. Right panel displays settings for obtaining FRET signal. Green column in graph indicates fluorescence of only mNeongreen in the red channel, i.e background signal. Red bar is red fluorescence and yellow is fluorescence in FRET settings. Pink bar is FRET fluorescence minus background fluorescence, actual FRET signal. Labels on columns indicate which fluorescent protein analyzed at which specific excitation/emission settings.

Task Three:

For this task we developed a microfluidics system with a local company and using various treatment methods and testing flow dynamics for the new chip we were able to determine a working combination to allow for long term growth of the pneumococcus. Figure 7 shows the first two hours of an overnight experiment to demonstrate the ability to induce the expression of labeled transporter machinery, which is indicated by a green signal.

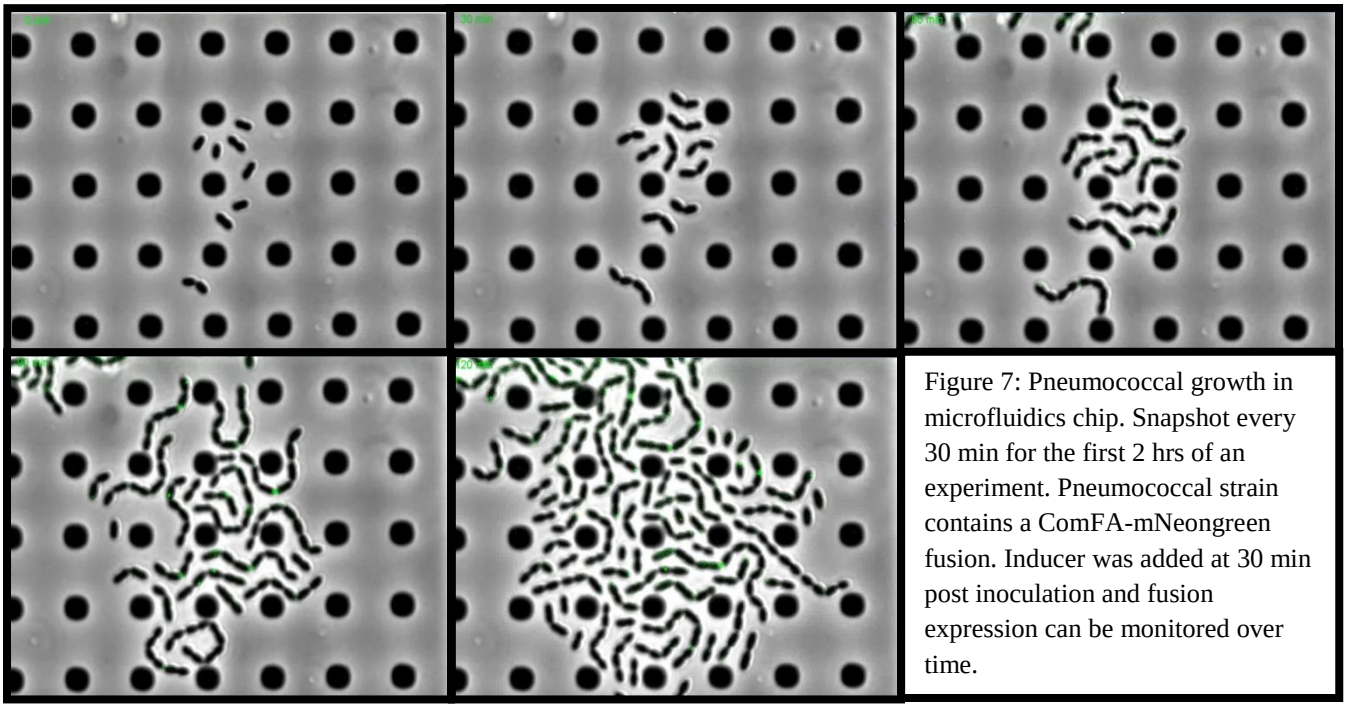

Figure 7: Pneumococcal growth in microfluidics chip. Snapshot every 30 min for the first 2 hrs of an experiment. Pneumococcal strain contains a ComFA-mNeongreen fusion. Inducer was added at 30 min post inoculation and fusion expression can be monitored over time.

To visualize DNA interacting with pneumococcal proteins, tagged ComEA was expressed in an endA mutant. ComEA is responsible for binding DNA, while EndA degrades a single strand of the DNA. With an EndA deletion mutant the introduced DNA is not degraded allowing us to visualize the DNA for longer periods. This allows us to test the best stains with our fluorescent proteins without degradation of the DNA. Various DNA stains and labeling methods were used to test the best method for detecting interaction with the competence machinery. Labeling methods include common DNA stains DAPI and propidium iodine, a modified dUTP (ThermoFisher, R1091) that incorporates into DNA during amplification followed by conjugation to a desired fluorescent protein (ThermoFisher, 62266), and the universal linkage system (ULS) PlatinumBright (Kreatech, GLK-004). As seen in Figure 8 three of the four labeling methods were successful in observing DNA-protein interaction. Propidium iodine worked well for labeling, but not DAPI. The modified dUTP incorporation produces signal, but not as intense as the propidium iodine. The ULS system seems to work the best for labeling as DNA localization was noted on most cells, but due to technical difficulties the protein fluorescence was not produced during this experiment.

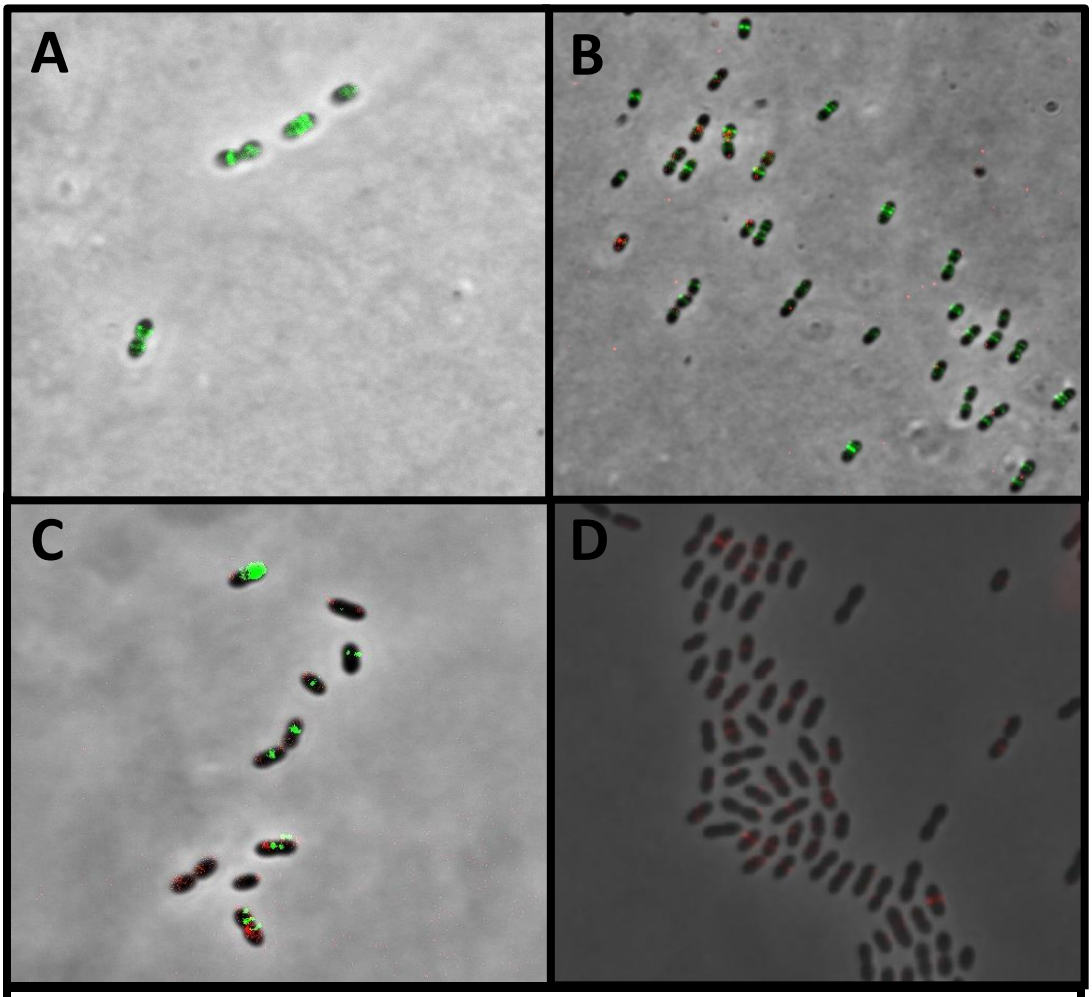

Figure 8: Different DNA labeling methods and interactions with ComEA. A) DNA stained with DAPI and ComEA labeled with mNeongreen. B) DNA stained with propidium iodine (red), ComEA-mNeongreen (green) and protein-DNA interaction (yellow). C) DNA labeled with modified dUTP and Dylight 650 (red) and ComEA-mNeongreen (green). D) DNA labeled with PlatinumBright (red).

Task Four:

This task was designed for optimization of fluorescent proteins and DNA labeling to determine what combinations produce the best FRET signal. We also tried optimization of fluorescent proteins by using a novel codon optimization method, which allowed a significant increase in fluorescent intensity for the tested mNeongreen protein. As shown in Figure 9 and 10 this optimization increases the fluorescent signal compared to the wildtype sequence for mNeonGreen, but not YFP.

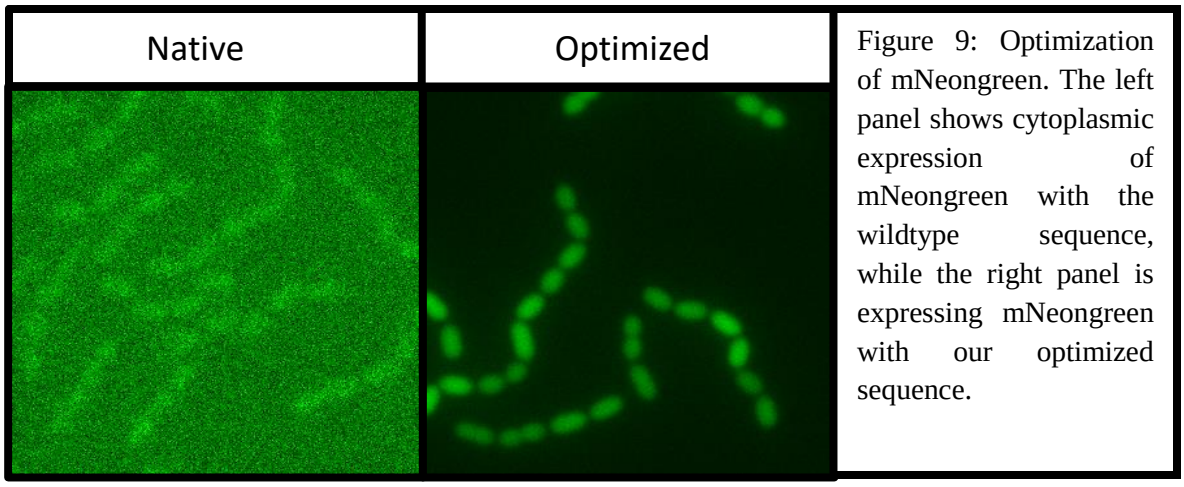

Figure 9: Optimization of mNeongreen. The left panel shows cytoplasmic expression of mNeongreen with the wildtype sequence, while the right panel is expressing mNeongreen with our optimized sequence.

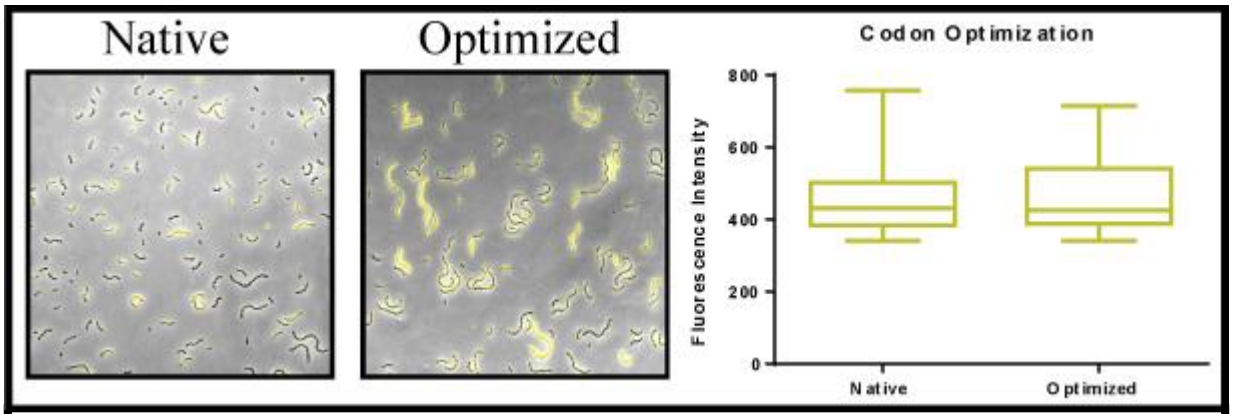

Figure 10: Codon optimization of YFP. Representative pictures of cytoplasmic expression of YFP with native sequence or with optimized sequence. Comparison of fluorescence intensity between tested YFP indicates that there is no increase in fluorescence based on codon usage.

Using codon optimized pairs of known FRET combinations we can test functionality for use within *S. pneumoniae*. Common FRET pairs include blue and yellow combinations as well as green and red, there is also evidence that mNeongreen and mTurquoise2 make a viable FRET pair. Using our codon optimization technique we have created fusions between mNeongreen and GFP with mScarlet-I, along with YFP and mNeongreen with mTurquoise 2 (Figure 11). These four different combinations also contain a TEV protease linker between the two to allow dissolution of the direct interaction between the two proteins. This allows us to determine FRET signal when the proteins are in close contact and then breaking that contact to visualize loss of FRET.

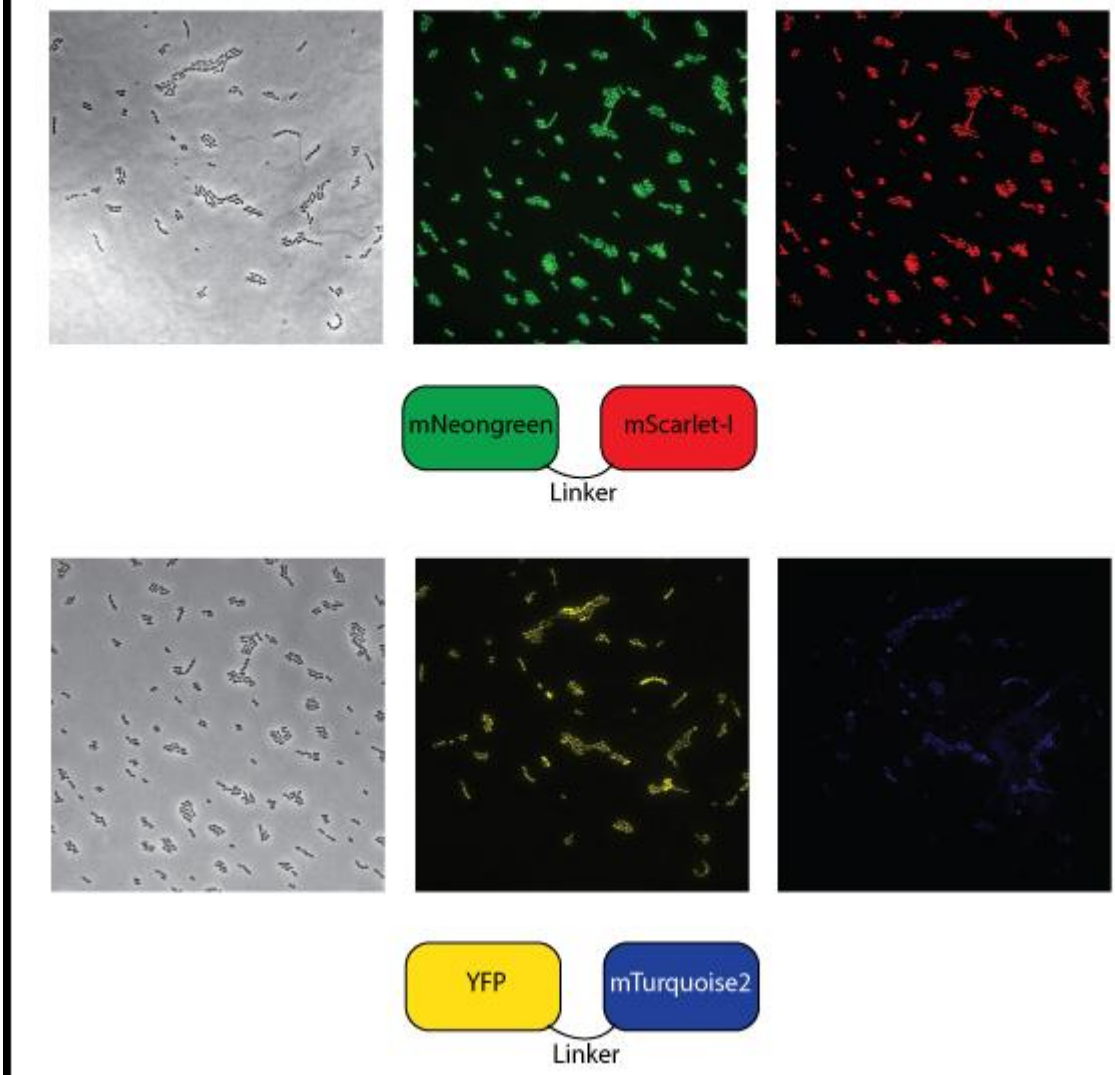

Figure 11: Characterization of optimized fluorescent proteins for use as a FRET pair. Cytoplamiscally expressed fluorescent proteins are linked by a cleavable peptide linker. Classic FRET combinations red-green and yellow-blue were created with a novel optimization program of the most photo-stable fluorescent proteins available. The fused proteins are able to be expressed and produce a bright signal that can be used to determine FRET compatibility.

Task Five:

This task was designed to measure FRET signal between adjacent fluorescent proteins and then to visualize multiple fluorescent foci to determine change in FRET signal when DNA is being imported. As seen in Figure 12 an intramolecular FRET between linked mNeongreen and mScarlet-I produces a viable FRET signal. This is obtained by exciting the mNeongreen at 470nm and reading output through a 605nm filter, which removes the majority of mNeongreen

signal from the mScarlet-I output. This is corrected for by removing background signal of mNeongreen through the same settings. After this is done the first panel between fused proteins (intramolecular) has an almost 20% increase in signal over background. Through visualization of mNeongreen fused ComFA and DNA labeled with a substrate that has similar fluorescent properties as mScarlet-I we can determine the FRET signal between protein and genomic DNA. FRET signal between these adjacent molecules produces about a 10% increase over background signal as shown in Figure 12.

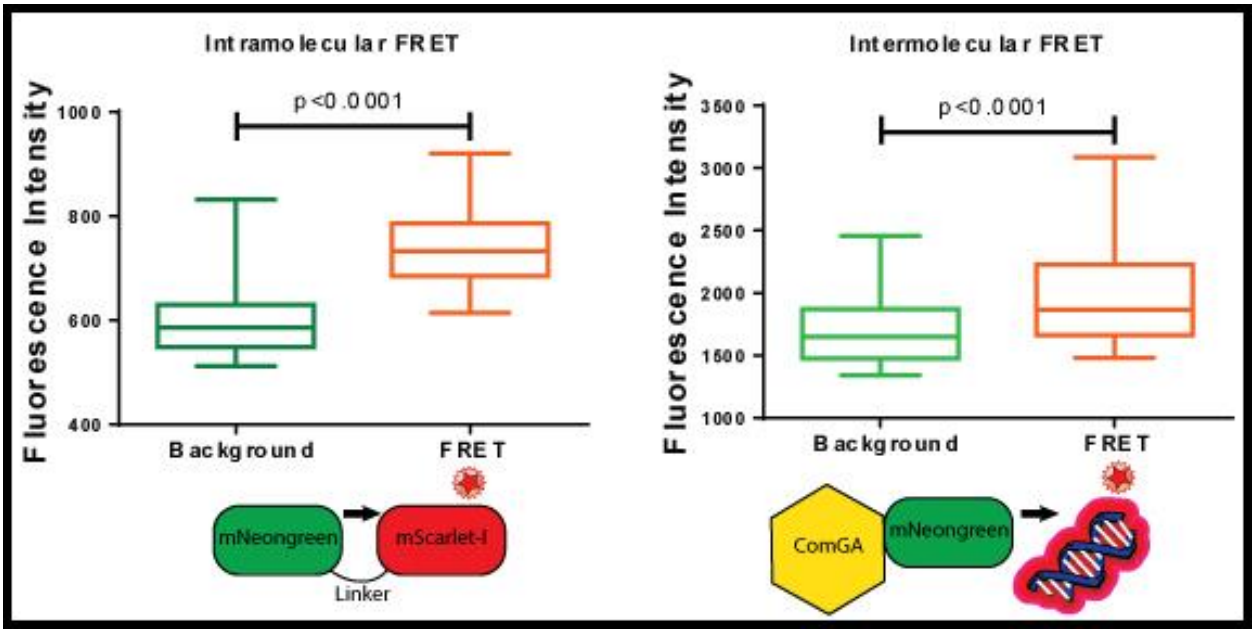

Figure 12: Recorded FRET signal between adjacent molecules. Fluorescence intensity was measured for cytoplasmically expressed mNeongreen excited with 470nm light and recorded through a 605nm bandpass filter (Background). The same settings were used to measure fluorescence intensity of cytoplasmically expressed mNeongreen fused to mScarlet-I. producing intramolecular FRET signal. Similar settings were used to observed background signal of ComGA fused with mNeongreen. This same construct was combined with fluorescently labeled DNA and the FRET signal between the labeled protein and the labeled DNA was recorded.

Through the same technique single foci can be tracked overtime with rapid exposure to determine a baseline FRET. Due to bleaching effects the fluorescence intensity reduces over time, but is still detectable after 300 exposures for some foci. At maximal DNA uptake by the pneumococcus this would result in 3,000 bp of DNA being imported over the 10 seconds of the experiment. Through altering exposure conditions, the number of activations can be increased from the current 300. Also more photostable fluorescent proteins are being developed which can increase the amount of exposure a fluorescent protein can endure before bleaching. In Figure 13 over a hundred foci are tracked over the indicated time. Each foci is standardized to 100% fluorescence to the first exposure and signal overtime is standardized to the first exposure strength and all foci values are averaged per time point. As expected there is a reduction in

strength over time, but this base line data can be used to determine if during DNA importation there are alterations in the FRET signal. Using this method and known DNA we can then start to track whether specific alterations in the signal correspond to certain nucleotides. Current exposure settings are a log slower than optimal DNA importation by pneumococcus but can be optimized for more rapid acquisition.

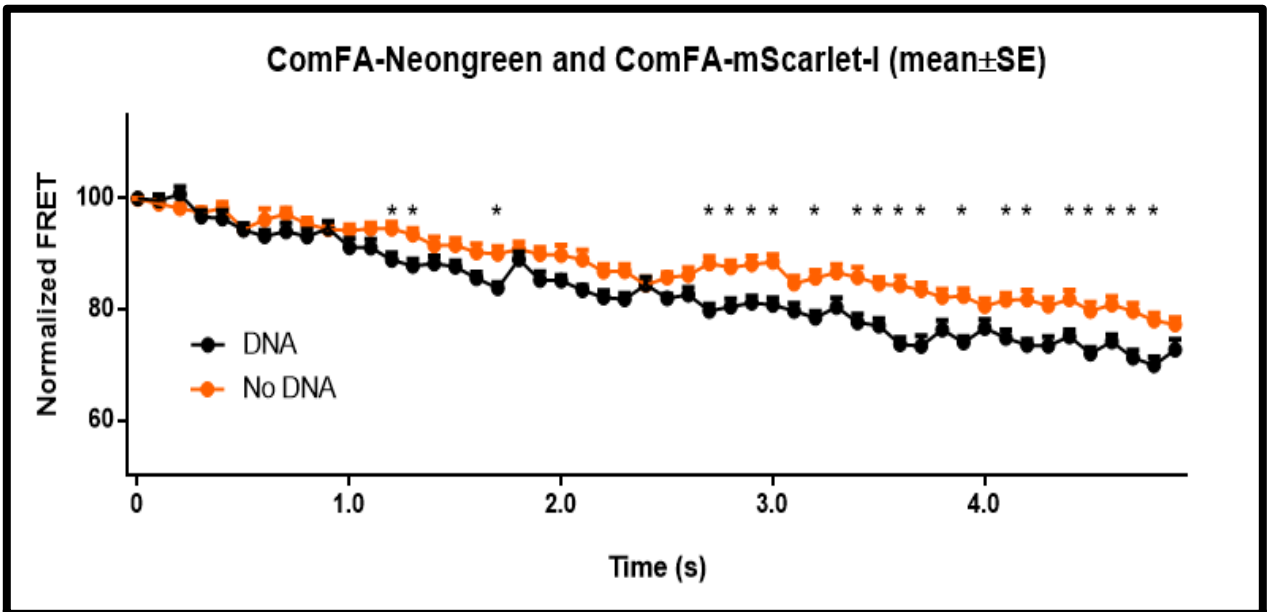

Figure 13: FRET signal of ComFA over time. Adjacent ComFA proteins labeled with either mNeongreen or mScarlet-I were tracked over a ten second period with excitation every 100ms. FRET signal for each foci was normalized to 100 and all ComFA foci FRET signals were averaged for each timepoint. The same strain was then exposed to exogenous genomic DNA and FRET signal was obtained and compared to FRET signal without DNA addition. Mann-Whitney test determined that indicated time-point were significantly different after addition of DNA. p-value < 0.05.

Task Six:

The final task was to utilize the previous data collected and test the ability to interpret the difference in FRET signal and how specific changes can be associated with incoming nucleotides. Using previous data about compatible FRET combinations the competence protein was labeled with mNeongreen and mScarlet-I and transformed into the same strain. ComFA was chosen because of the intensity of its signal allowing for the best FRET signal possible. As seen in Figure 13 we were able to obtain a FRET signal with and without the addition of DNA and that specific timepoints displayed significant differences when DNA was added. We further wanted to test these observations by altering the type of DNA being imported and determine if the signal acquired altered based incoming DNA. Previous research has shown that the maximal rate of DNA importation by the pneumococcus is one nucleotide every 10 milliseconds. Due to this we explored FRET signal at multiple read rates.

First, we tested the signal acquired when only the donor fluorophore is present, mNeongreen in this experiment. This is done to verify that the alterations of our signal acquired during our FRET experiment is due to changes based on the DNA being imported. Figure 14 shows that when only ComFA fused with mNeongreen is expressed the signal acquired does not significantly differ based if DNA is added or what type of DNA is added.

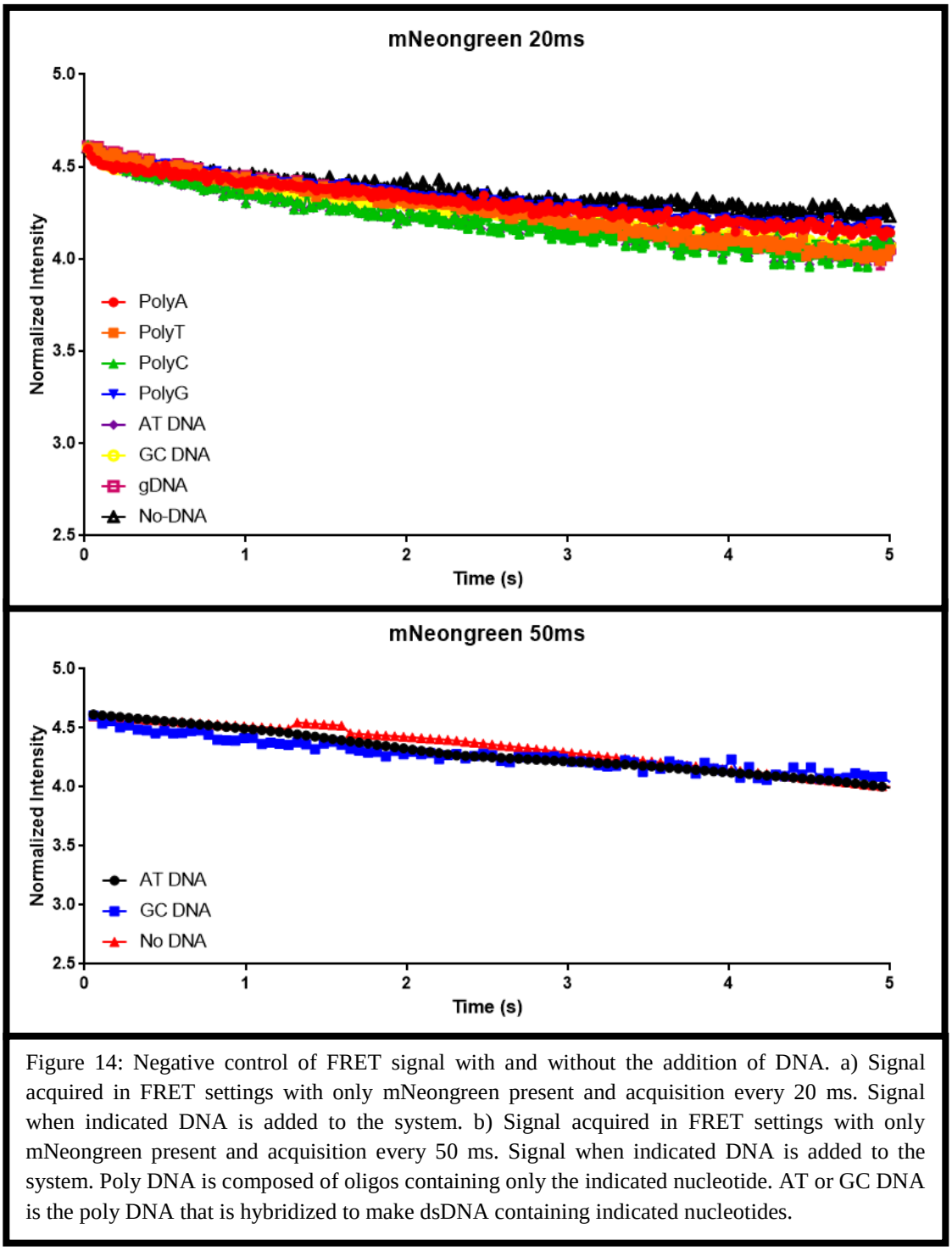

Figure 14: Negative control of FRET signal with and without the addition of DNA. a) Signal acquired in FRET settings with only mNeongreen present and acquisition every 20 ms. Signal when indicated DNA is added to the system. b) Signal acquired in FRET settings with only mNeongreen present and acquisition every 50 ms. Signal when indicated DNA is added to the system. Poly DNA is composed of oligos containing only the indicated nucleotide. AT or GC DNA is the poly DNA that is hybridized to make dsDNA containing indicated nucleotides.

We now know that any background signal or perturbations in signal observed during DNA importation is likely due to the type of DNA being imported and not other effects. With this information we can now test if there are difference in FRET signal when DNA is imported. As seen in Figure 15, regardless of the time of signal acquisition there is an observed difference in FRET signal when DNA in imported. Individual values at each time point is the average intensity of several hundred points. Using this information, we can see that there is an average difference in FRET signal that is dependent upon the DNA added and the speed of acquisition is independent of the signal difference. There are variations in the signal based on acquisition time, but the general trend that DNA importation alters the FRET signal is consistent. Future work could focus on single foci and individual changes through time. This will help map specific changes due to certain nucleotide importation.

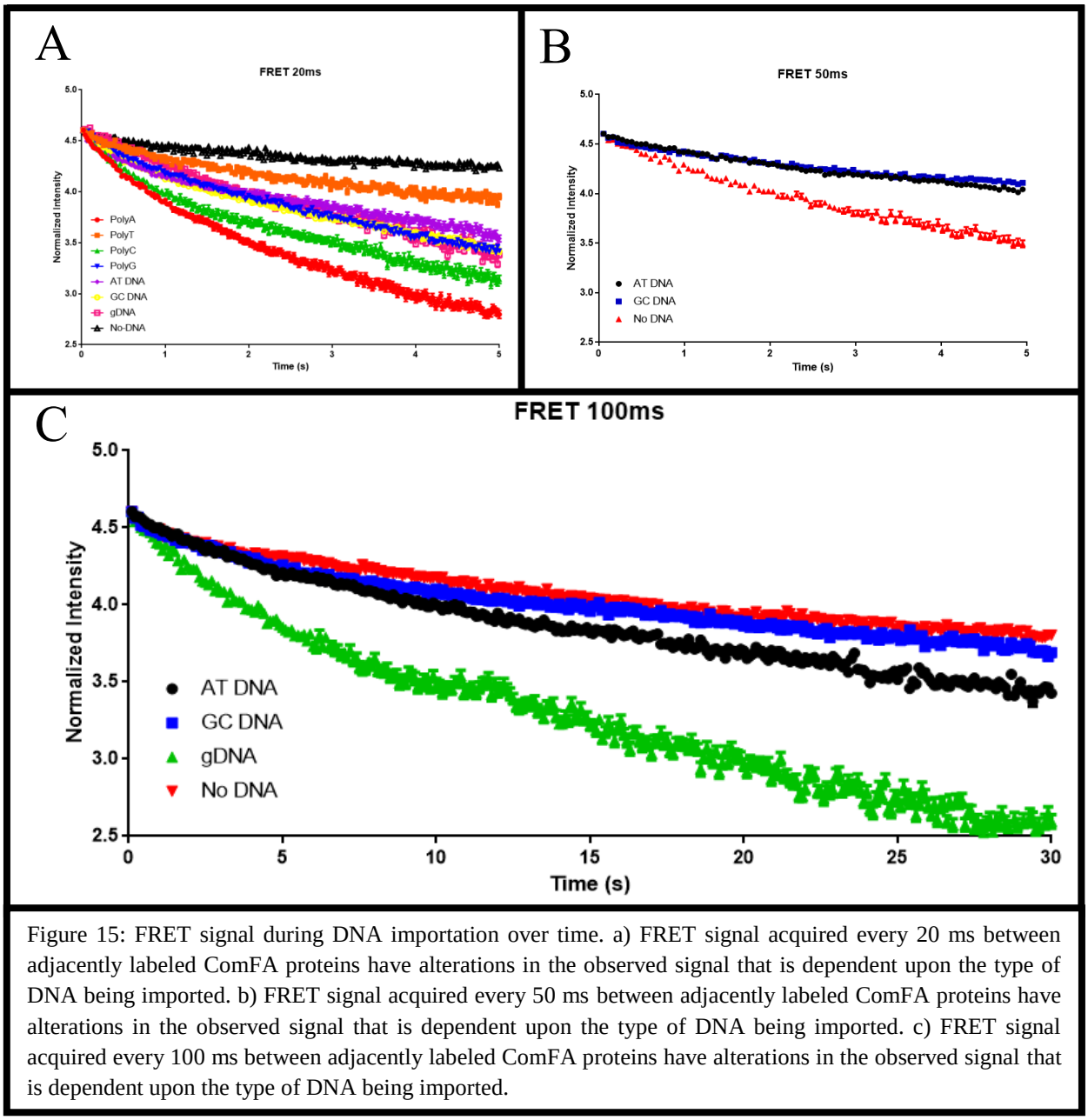

Figure 15: FRET signal during DNA importation over time. a) FRET signal acquired every 20 ms between adjacently labeled ComFA proteins have alterations in the observed signal that is dependent upon the type of DNA being imported. b) FRET signal acquired every 50 ms between adjacently labeled ComFA proteins have alterations in the observed signal that is dependent upon the type of DNA being imported. c) FRET signal acquired every 100 ms between adjacently labeled ComFA proteins have alterations in the observed signal that is dependent upon the type of DNA being imported.

We now know that any background signal or perturbations in signal observed during DNA importation is likely due to the type of DNA being imported and not other effects. With this information we can now test if there are difference in FRET signal when DNA is imported. As seen in Figure 6, regardless of the time of signal acquisition there is an observed difference in FRET signal when DNA in imported. Individual values at each time point is the average intensity of several hundred points. Using this information, we can see that there is an average difference in FRET signal that is dependent upon the DNA added and the speed of acquisition is independent of the signal difference. There are variations in the signal based on acquisition time, but the general trend that DNA importation alters the FRET signal is consistent. Future work could focus on single foci and individual changes through time. This will help map specific changes due to certain nucleotide importation.

Summary of Findings:

As demonstrated throughout this report great progress has been completed on all task and a general understanding of the viability of this method for DNA sequencing has been established. Through the work done here it is hypothesized that this method can be used to sequencing DNA being imported through the pneumococcal DNA transport machinery. While through the year of research done on this project was not able to sequence de novo DNA and obtain a sequence it demonstrates the viability of this method. Through the creation of multiple fusion vectors, DNA labeling methods, and FRET pair optimization we can detect a measurable difference of the FRET signal when DNA is imported. The data demonstrated in the final figures is the average FRET signal over hundreds of fluorescent foci and indicates that there is a difference. Future work could identify the difference between individual foci and the average background to determine if the amplitude change corresponds to a specific nucleotide. Much work can be done on this project for optimization of FRET output, altering DNA binding protein ComEA to allow for more efficient ssDNA or RNA binding, and the possibility of incorporating this as a cell free system.

Table 1:

Task	Milestone/Deliverable	Start Date	Completion Date	Percent Complete	Figure
1.1	Generate unencapsulated and pneumolysin negative *S. pneumoniae* strain for biosafety level 1 work	9/1/2017	9/8/2017	100	
1.2	Using fluorescent protein mNeongreen create C terminal and N terminal fusions of multiple competence transporter machinery	9/8/2017	10/13/2017	100	1, 2
1.3	Delete individual genes from chromosome that are being tagged with fluorescent marker	10/13/2017	11/3/2017	100	
1.4	Under full induction determine ability of individual terminal fusions to transform using standard	11/3/2017	11/17/2017	100	3

	transformation procedures and incorporation of resistance marker				
1.5	Induce expression of individual fusions at levels known to produce ~1-5% expression to obtain single functional DNA importer per strain	11/17/2017	11/24/17	100	4
2.1	Various fluorescent proteins will be tagged to DNA transporter machinery found to be functional from task 1 and functionality will be retested	11/24/2017	1/5/2018	100	5
2.2	Revised: Determine DNA transporter components that interact closely enough to produce FRET signal	1/5/2018	1/26/2018	100	6, 9, 10
2.3	Revised: Obtain FRET signal between interacting proteins	1/26/2018	2/2/2018	100	6
3.1	Optimize microfluidics setup for our state-of-the-art Leica TIRF microscope and flow settings to allow for optimal DNA transport and cell viability	2/2/2018	2/16/2018	100	7
3.2	Label test genomic DNA and visualize DNA import using TIRF microscopy	2/16/2018	3/13/2018	100	8
4.1	Use various DNA dyes for best signal with mNeonGreen tagged proteins	4/13/2018	5/3/2018	100	8
4.2	Alter tagged transporter components with other fluorescent proteins and determine FRET interactions with relevant available DNA labels.	5/3/2018	5/24/2018	100	11
5.1	With previous information about labeled machinery that produces a viable signal use a combination of labeled transport machinery with single labeled DNA to obtain different signals based on incoming DNA	5/24/2018	6/8/2018	100	12
5.2	Using transport machinery labeled with multiple fluorescent proteins measure FRET signal of imported DNA that has multiple fluorescent signals	6/8/2018	6/22/2018	100	13
6.1	Synthesize DNA that contains long stretches of only single nucleotides and read signal during import to obtain baseline signal for importation of specific nucleotides	6/22/2018	7/19/2018	100	14, 15
6.2	Synthesize DNA that contains stretches of all nucleotides in a single strand to obtain data of FRET signal as importation switches from one nucleotide to the next	7/19/2018	8/30/2018	100	14, 15