Enabling Sample Tracking in and Reducing Variability in DNA Library Construction

By
David A Penake
B.S. Materials Science and Engineering, Northwestern University (2001)

Submitted to the Sloan School of Management and the Department of Materials Science and Engineering in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Materials Science and Engineering

In Conjunction with the Leaders for Manufacturing Program at the
Massachusetts Institute of Technology
June 2006

©2006 Massachusetts Institute of Technology. All rights reserved.
This page has been intentionally left blank.
Enabling Sample Tracking in and Reducing Variability in DNA Library Construction

By
David A. Penake

Submitted to the Sloan School of Management and Department of Materials Science and Engineering on May 12, 2006 in partial fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Materials Science and Engineering

Abstract

The Broad Institute is a world leader in genomic sequencing. The Institute gained prominence during the Human Genome Project and has continued to be the nation’s leading public laboratory for genomic sequencing projects. The Broad Institute has done this by taking advantage of economies of scale, rapid process innovation, and new process implementation. Strong execution has made The Broad Institute the lowest cost and highest quality sequencing operation in the country.

As biological science evolves, The Broad Institute will focus on low-volume, highly targeted areas of a genome rather than high-volume, full genome sequencing projects. In order for the Institute to maintain high-quality output at low cost, it must change its operations strategy from mass production to flexible, lean production.

The operations organization has begun the lean production journey by hiring operations focused managers and initiating a Six Sigma program to utilize their employees to innovate and optimize processes. This effort has been supplemented with 5S (systematic organization) and information systems that improve communication between groups.

This thesis will focus on engineering and management issues relating to The Broad Institute and its quest to maintain quality and cost leadership among genomic sequencing laboratories. Specifically, the thesis will focus on the following items:

- Preventing organism sample swaps by implementing operations changes that minimize operator error and provide electronic sample tracking through the use of Datamatrix (two dimensional) barcode technology.

- Minimizing inherent process variability by increasing reagent inventory turnover and eliminating the use of expired chemicals.

- Increasing the flexibility of the production process to easily handle high-mix, low-volume projects, such as the Cancer Human Genome Project.

Thesis Supervisor: Roy Welsch
Title: Professor of Statistics and Management Science

Thesis Supervisor: Darrell J. Irvine
Title: Assistant Professor of Materials Science and Biological Engineering
This page has been intentionally left blank.
I wish to acknowledge the Leaders for Manufacturing Program at MIT for its support of this work.

I would also like to thank Robert Nicol, Marcia Lara, and the Molecular Biology Production Group at the Broad Institute of MIT and Harvard.

I would like to thank all members of the Class of 2006 LFM Fellows for their help and friendship over the last two years.

I want to give special thanks to my parents for making education and hard work important aspects of my life. I appreciate all of your sacrifices that have provided me such wonderful opportunities.

Additionally, I would like to thank Mike, Michele, the White’s, all my family and friends for their support and guidance throughout my life. None of this would have been achievable without your numerous sacrifices and continuous support.

Finally, I want to thank my beautiful fiancée, Adrienne. You make my life special each and every day. I am always amazed by you and understand the sacrifices that you have made to allow me to attend MIT. Thank you so very much for your love, your advice, and your constant support.
This page has been intentionally left blank.
Table of Contents

1. INTRODUCTION
   1.1 Background
   1.2 Challenges to the Future of Genome Sequencing
   1.3 Strategy to Maintain Quality and Cost Leadership
   1.4 Conclusion

2. STATEMENT OF THE PROBLEM

3. APPLICATION OF LEAN MANUFACTURING TECHNIQUES
   3.1 Preserving Quality in DNA Transformation
   3.2 Analysis and Improvement Methodologies
   3.3 Transformation Process Mapping
   3.4 Highlighting Risky Processes in Transformation
   3.5 Re-engineering the Transformation Process
   3.6 Implementation Problems
   3.7 Outcomes of Project Implementation
   3.8 Conclusion

4. CONTROLLING VARIABILITY IN THE MOLECULAR BIOLOGY PRODUCTION GROUP
   4.1 Technical Background – Enzymes and Library Construction
   4.2 Preventing Material Degradation
   4.3 Understanding the Existence of Variation
   4.4 Reducing Variation in Materials Supply
   4.5 Future Areas of Focus to Reduce Variation
   4.6 Conclusion

5. STRUCTURED PROBLEM SOLVING
   5.1 Prior Outcomes of Quality Issues
   5.2 Quality Problems in Molecular Biology
   5.3 Outcomes from Structured Problem Solving
   5.4 Conclusion

6. ORGANIZATIONAL STRUCTURE AND EMPLOYEE DEVELOPMENT
   6.1 The Broad Institute – A Growing Organization
   6.2 Six Sigma Training
   6.3 Implementation of Structured Processes
   6.4 Conclusion
7. **SUSTAINING IMPROVEMENTS**

7.1 Sustaining 5S  
7.2 Extending Sample Tracking  
7.3 Maintaining a Learning Organization  
7.4 Conclusion

8. **CONCLUSIONS**

8.1 Developing Lean Manufacturing Capabilities  
8.2 Educating Employees about Organizational Goals  
8.3 Final Recommendations

9. **AUTHOR’S BIOGRAPHY**

10. **REFERENCE**
1. Introduction

The Broad Institute is a unique organization. It first gained fame during the Human Genome Project. The project was one of three scientific superprojects; Apollo, the Manhattan Project, and the Human Genome Project, that have taken place in the United States.

During the project, The Broad Institute, formerly the Whitehead Institute Center for Genome Research, was the primary laboratory for sequencing operations. During that time, the Institute sequenced nearly 40% of the genome, or 1.2 billion base pairs\(^1\), helping the public sector meet the goal of mapping the Human DNA by 2003.

Since 2003, the Institute has completed other genomes including; mouse, cat, dog, cow, elephant, armadillo, rabbit, tenrec, and guinea pig\(^2\), each having their own purpose for helping to improve our understanding of science and the building blocks of life, DNA. The mission of the Institute also changed with a large donation from Eli and Edythe Broad in 2003.

In May 2004, the new mission of the Institute was announced;  
*The Broad Institute is a research collaboration of MIT, Harvard and its affiliated Hospitals, and the Whitehead Institute, created to bring the power of genomics to medicine*\(^3\).

This mission has altered the future of the Broad Institute and the types of work that it will be completing. Future projects for the center will likely include significant portions of the Human Cancer Genome Project, a project that will involve sequencing portions of the DNA of more than 50\(^4\) cancer types. This project, and others like it, are a fundamental change from large scale mammalian projects and will require changes to the sequencing operations strategy, focusing efforts on more projects with greater flexibility.

1.1 Background

Genome sequencing operations consist of five primary groups; Materials, Molecular Biology Production, Core Sequencing, Detection, and Finishing. Each group performs a critical role in transforming the raw material of deoxyribonucleic acid (DNA) into data that can be used for scientific application.

At a high-level, the genome sequencing process is a method of detecting the coded sequence within DNA. The code is made up of four base-pairs that make up the entire genome. These base pairs are adenine (A), thymine (T), cytosine (C), and guanine (G). In the double helix formation of DNA, A always bonds with T and C always bonds with G. Therefore, we can decode one of the helix structures and transpose the information to decode the other. Figure 1\(^5\) depicts an artistic view of DNA, the double helix structure, how it relates to chromosomes and cells within the body.
These letters make up all the genes, proteins, and amino acids that regulate the functions of the organism associated with DNA. To give a sense of perspective, the Human DNA is made up of 3 billion base pairs and about 20,000-25,000 genes. Despite the size of the genome, nearly all encoded DNA is identical for the human species. People have small mutations, or differences, between one another, but the majority of the information is the same for all people. This similarity allows scientists to use the information to characterize populations and diseases to find cures that may relate to many people rather than a specific population.

The operations at the Broad Institute use the Whole-Genome Shotgun (WGS) method of sequencing. Below is a brief summary of each department and how they contribute to successfully sequencing DNA and publishing the information for use by the public and researchers.

1.1.1 Materials
The materials supply group provides all reagents, disposables, and quality inspections for groups at the Broad Institute. Their processes include material delivery of reagents that are used in standard production processes, production of in-house reagents, materials required for genome sequencing, and qualification of in-house and outsourced materials or reagents. The group maintains supplier relationships and provides employees at the Institute with the materials that are required for their jobs at the time they need them.

The materials group uses an ERP system to manage inventory and materials usage. In addition, they attempt to minimize inventories by forecasting production needs and ordering materials and quantities to support the operations. Due to the high variation in
levels of production sequencing, forecasting has proven difficult and the group is attempting to improve methods of gathering information and ordering materials in a timely manner.

The group is the last barrier of protection against poor quality material. The group conducts quality control on materials that are supplied from vendors as well as those created in house. These quality controls are necessary and time consuming, but prevent most costly quality errors by catching the problems before they are used in production.

1.1.2 Molecular Biology Production
The Molecular Biology Production Group (MBPG) is the first group to begin DNA preparation for sequencing. At this point, a full strand of DNA is fed into a machine that shears the DNA using mechanical force. This shearing creates DNA strands of varying length (4,000 – 40,000 base pairs) depending on the application and production process. After shearing, the DNA is inspected for quality, at which point, the technician removes the strands that are not the desired length.

DNA vector ligation is the next step in the process. The ligation process requires the molecular biologist to trim the ends of DNA strands and create a standard base pair sequence at each of the ends to allow for easy insertion into another cells’ DNA, as depicted in Figure 2. In addition to standardizing the ends, the ligation step does a finer job of sizing the strands so that the great majority of the strands are of the desired length with a standard deviation of 100 base pairs or less. These strands are then inserted into a plasmid, or circular DNA strand. This process is illustrated in Figure 3. Once the concentrated DNA ligation solution has been created, the DNA is ready to be amplified through biological processes.

![Figure 2. Ligation reaction creating standard ends of target DNA.](image-url)

The DNA is then diluted before it is used in high volume production. This is done because each ligation has the ability to be split into 30 – 50 dilution sets based on the quality and efficiency of the ligation. Therefore, one week of labor for a ligator can produce two or more weeks of work for the transformation process.
Once diluted, the transformers add the DNA to cells through a process called electroporation. During electroporation, a machine shocks a solution of DNA dilution and electro competent cells. The shock causes the cell membrane to loosen so that the cells can take on the target DNA plasmid as their own.

**Inserting a DNA Sample into a Plasmid**

At this point, the DNA strand that is desired to be sequenced has been implanted into a living cell that now has DNA that includes it’s own, plus the desired strand. The solution of cells and a few other chemicals are then spread onto large plates coated with a growth medium and incubated for seventeen hours. This growth period allows us to amplify the strand of DNA that we have created because the cells will divide many, many times during the seventeen hours and each time they divide, the cell replicates the DNA to create another cell.

The plates are checked to ensure that a desired number of cell colonies (1,500 – 3,000) have been created and that the material is contamination free before it is sent downstream to Core Sequencing.

### 1.1.3 Core Sequencing

The Core Sequencing group receives the colonized plates from MBPG and uses machines to transfer the colonies from the large, agar plates to smaller plates with 384 wells. Each colony that is chosen resides in its own well and is prepared for further sequencing operations in a downstream process.
Within Core Sequencing, there are numerous processes that remove the DNA from the *e. coli* cells, kill the *e. coli* cells, amplify the remaining DNA and tag the last base pair of the DNA strands with a fluorescent dye that allows the detection machines to note which of the four bases are on the end of the string. These processes are highly automated and use machines to handle the materials and transfer samples from one to multiple plates as the DNA continues to amplify throughout the process.

Core Sequencing is also the primary material usage cost driver. Two materials that the group uses, Big Dye and Temple-Phi, drive more than 60% of all material consumption costs at the Institute and are constantly being investigated to reduce variation from sample delivery and improve the efficiency with which materials are used. Figure 4 depicts the consumable operations cost breakdown\(^8\).

**DNA Sequencing Consumable Cost Breakdown**

![Pie chart showing DNA Sequencing Consumable Cost Breakdown]

**Figure 4. Broad Institute consumable cost breakdown.**

Core Sequencing produces DNA that has been prepared for use in the detection machines. Through the reactions that occur during the process, the DNA strands are now marked with a dye to be read at the laser and have significantly different lengths based on the reaction kinetics. This will allow the detection machines to order the base pairs correctly during processing.

1.1.4 Detection

The detection process is where the DNA material is transformed into data that is useful for research. Detection is the last production process step and the capacity constraint of the factory.

The Broad Institute has 117 high throughput genome sequencing detectors. These machines drive the capital cost of the Institute at ~$350,000 each and limit the throughput of all operations to roughly 1.2 gigabase pairs per week despite running on a 24 hour per day, 7 day per week schedule.
The detectors transform the material into data through the use of capillary action. Inside of the detectors are many small tubes that guide DNA from the plates to the detection laser. Through this process, shorter stands of DNA move more quickly than longer strands. Once they reach the end of the tube, there is a laser that excites the dye on the end of the DNA strand and records which base the dye is associated with. As all of the lengths of strands are detected, a file is created that recreates the coded sequence of a full length of the strand.

The detection data is very fragmented and only provides the code of a small portion of the genome (~0.0006% for humans). This data must be compiled and placed in the correct order for the information to be useful.

1.1.5 Finishing
The finishers are a group of people that take the data from the detectors and assemble the genome. They use computer algorithms to help them place the data into the genome at the correct spot. In addition, the finishers must fill the gaps in the genome that are created when the snippets do not line up perfectly. In order to do this, they must run experiments that reveal the hidden genomic data. If successful, they can fill the gaps quickly. If it is a more difficult part of the genome, the finishers must order new ligations to be created for a very specific part of the genome. If this happens, the DNA goes through the entire production process a second time.

Once high quality data has been produced, the finishers must upload the data to the internet for free use by the public. This data is used by research institutions and companies worldwide to gain a better understanding of life and the biological processes that govern their existence.

1.2 Challenges to the future of genome sequencing
The initial design of the genome sequencing center was based on serial production of large, single genomes. The center was created to produce high-quality, low cost-per-read sequencing data to complete the map of the human genome.

The Broad Institute quickly became the leader in both cost and quality among laboratories in the National Human Genome Research Initiative (NHGRI). Through economies of scale, innovation, and strong execution, the Institute has maintained the lead in both cost and quality.

During the Human Genome Project, innovation led to major breakthroughs in process efficiency which led to decreasing costs associated with genome sequencing. As Rob Nicol illustrates in Figure 5, these costs have led down a path similar to that associated with Moore’s Law for semiconductor costs per transistor. Although these breakthroughs led to massive improvements in the cost of sequencing a genome, they were rarely created on flexible platforms.
After the completion of the Human Genome Project, the Broad Institute began to earn grants based on smaller projects that required sequencing of partial or multiple organisms. These new grants required flexibility that was not in place since operations had been designed around mass production. Unfortunately, these changes have caused quality and cost issues to rise with respect to many aspects of sample tracking and cost associated with materials and inventory.

To combat this problem and build flexible, operational capabilities, the Broad Institute hired several operations focused managers into high-level positions and allowed them flexibility to change the organization. The managers have focused on implementing programs that are necessary to maintain quality and cost leadership against other NHGRI funded laboratories while building flexibility into the production system.

1.3 Strategy to maintain quality and cost leadership

The Broad Institute has a strategy to maintain leadership among other laboratories through:

- Rapid process innovation and implementation.
- Use of cutting-edge, high-efficiency sequencing techniques.
- Operational execution to minimize cost and process variability.

Rapid process innovation and implementation

The Institute changes their entire Core Sequencing operations platform every twelve months. During process revisions, old equipment is removed and new equipment is installed and quickly qualified for production. In between the implementation periods, the process development group is creating machines that will work more efficiently with new sequencing processes and platforms that will be used in the future.

Use of cutting-edge, high-efficiency sequencing techniques
The Broad Institute must always be using the best sequencing technology to maintain leadership against other sequencing centers. In order to do this, the process development group is also constantly testing prototypes of detectors that have been created to provide the next generation of DNA sequencing. This development process allows the Institute to have insight to maximize capacity and create a robust process.

In addition to testing prototypes, the Institute is often heavily involved in helping the manufacturer fine tune their equipment and will set themselves up for priority ordering once the machine becomes commercially available. This allows them to have the first mover advantage over other sequencing centers and minimize the learning curve associated with process changeovers and revisions.

**Operational execution to minimize cost and process variability**

The Institute has hired managers into high-level positions with experience in industries outside of biology to create an outsiders perspective to the management of sequencing operations. The head of operations has focused on implementing quality programs within the Institute as well as bringing in people with similar industrial experience to provide a new perspective to operations.

The programs that have been initiated at the Broad Institute have included a Six Sigma program that many of the technicians take as part of their development plan. The Institute is using this program to maximize worker capabilities. The goal is to provide a standardized method to improve processes and innovate on the current platform. Additionally, this training will provide a standard method of problem solving that will minimize confusion between coworkers trying to solve a difficult process issue. Six Sigma is the largest current initiative, but the operations organization has also been exposed to the ideas of lean manufacturing, such as, 5S, operations redesign, and value stream mapping.

Along with these programs, the Institute has a history of hiring interns from the Leaders for Manufacturing Program at MIT Sloan. These interns are placed directly within a department in the organization and find a project that they believe will have a significant impact on the operational capability of the organization. The projects have introduced Broad employees to the concepts of 5S (systematic organization), inventory management, cycle time reduction, operational metrics, materials management, and work prioritization. Many of these internships have been considered highly successful and have had significant impacts on the organizations ability to complete projects on time and within budget.

In order to maintain cost and quality leadership, the Broad Institute must change its operations strategy from mass production to a lean, flexible system that will be able to handle multiple, low-volume projects flawlessly. The Institute has continued to invest in its operational capability to remain the NHGRI’s center of choice.

**1.4 Conclusion**
This thesis will describe and analyze methods that the Broad Institute is using to reach its strategic goals. Additionally, this paper will investigate the culture and organization of the Institute and provide recommendations for further success in genomic sequencing.
2. Statement of the Problem

The problem facing the Broad Institute is to build operational flexibility while continuing to lead competitive laboratories in quality and cost metrics. As the Institute worked toward this goal, the production profile began including many diverse projects on similar timelines. Six years ago, the Institute was only sequencing one genome, while current processes have ten or more genomes in production at all times. This change required that the Institute improve systems within the organization to prevent misregistration and maintain high quality. Unfortunately, those systems were not in place.

In 2004, the Broad Institute became aware that it had mistakenly published data listed under the wrong organism. Since the data had been published to the internet, researchers had been using the data to develop therapies and understand biological function. The affected researcher contacted the Broad Director, who immediately moved to correct the problem. In order to prevent this from happening again, and to reinforce the Broad Institute’s commitment to quality, Eric Lander, the Broad Director had the Institute create a program called Organism Checker.

*Organism Checker* takes data from the partial sequence that is read off of the detectors and compares the data to all publicly available genomic sequence data for that organism. After the comparison, it gives a value that tells the Institute how confident the system is that the data provided is actually from the organism that was specified. Generally, the system has very high confidence when the data is correct and significantly lower confidence when the data has been misregistered.

The system prevents incorrect data to be published to the internet, but it does not address the problem of misregistration that happens within the operations of the Institute. Additionally, the *Organism Checker* has several weaknesses;

- It cannot tell the Institute what organism the data actually came from.
- It will not be able to tell the difference between sequences of the same organism with slight mutations (i.e. human leukemia vs. human myeloma).
- Since the Broad Institute often publishes genomic sequence on an organism first, there may not be any data in public databases to check the sequence against.

Additionally, it is very difficult to transfer data from the wrong organism to the correct organism if the Institute is able to find the organism that the data came from. This means that most misregistration data must be disposed of and cannot be used for a project. Waste of sequence data inhibits the operations ability to meet cost targets and wastes constraint capacity.

The operations system within the MBPG has a high risk of quality errors due to the complexity of the process and the number of samples that are processed within the lab at any point in time. The type of quality error that has had the most significant impact on the organization in recent years is the opportunity for sample swap, or sample misregistration.
The risk of misregistration is currently high within the organization and is detailed in Table 1. The table also includes data that was compiled about risk after the project redesign. In addition to the high-risk associated with sample misregistration, there are a number of other factors have are detrimental to the area when a sample swap occurs. These factors include:

- A significant cost to society if incorrect data is published.
- Lost throughput at the constraint when data cannot be converted.
- Management troubleshooting of the problem without data necessary to arrive at root cause.
- Mistrust between transformers and management because blame is placed without root cause data.

### Redesign Improvement Data Summary

<table>
<thead>
<tr>
<th></th>
<th>Before Redesign</th>
<th>After Redesign</th>
<th>% Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood of ligation misregistration</td>
<td>0.080000%</td>
<td>0.00001%</td>
<td>13714286%</td>
</tr>
<tr>
<td>Plates / mixup</td>
<td>1250</td>
<td>171428571</td>
<td></td>
</tr>
<tr>
<td>Likelihood of transformation misregistration</td>
<td>0.002000%</td>
<td>0.000008%</td>
<td>25000%</td>
</tr>
<tr>
<td>Plates / mixup</td>
<td>50000</td>
<td>125000000</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Statistical analysis of misregistration risk within MBPG before and after the operations redesign.

Table 2 compiles displays the misregistration data from the Broad Institute operations. This includes the MBPG, as well as the other functional areas within the operations organization. As one can see, the hypothesized data shown in Table 1 matches closely to the actual data from Table 2. After full implementation of the transformation process, improvements will be able to be tracked against predicted outcomes.

### Organism Checker data May 2005 - December 2005

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th>Plates / misregistration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood of a single plate misregistration</td>
<td>0.89%</td>
<td>113</td>
</tr>
<tr>
<td>Likelihood of a large scale misregistration (&gt;5% of daily production)</td>
<td>0.0014%</td>
<td>73,243</td>
</tr>
</tbody>
</table>

Table 2. Broad Institute misregistration data.

The Organism Checker was put in place after a data swap had been highlighted by a researcher and reported their findings to Eric Lander and the funding agency. Once
Organism Checker was fully operational, the number of misregistrations that have been published are believed to be zero because this is a final quality measure that has been put in place. Unfortunately, the system has limitations and cannot tell the difference between different genomes that came from the same organism, such as different types of human cancers. Given this limitation, we cannot say with 100% confidence that no incorrect data has been published.

Despite the limitations of the quality system, it has done an excellent job of removing nearly all sample misregistrations from being published under the incorrect data set. Unfortunately, it is an end-of-line quality check that requires a great deal of work for the managers to investigate the cause of the problem and prevent it from happening in the future. In addition to the man-hours required for the investigation, management must make a decision about whether to move the data to the correct organism file. This is very difficult to do and if there is an error during placement, the data will be published incorrectly. Due to the risks associated with this technique, it is only used when the data is critical to the completion of the project and a deadline is approaching.

If the Institute is early in the project or has a great deal of extra data, the swapped sample data is discarded, which is wasted throughput of the constraint.

This thesis will focus on highlighting the area that is most likely to have a sample swap and use techniques of lean manufacturing to redesign the operations of the highlighted areas and put information technology systems in place to track production and provide the data necessary to find root cause of a sample swap, and prevent this from causing a problem with data integrity of the samples. Addressing these issues will:

- Improve sample tracking within The Broad Institute.
- Reduce the risk and cost associated with quality errors.
- Increase the availability of data for process characterization.
3. Application of Lean Manufacturing Techniques

Lean manufacturing techniques have proven useful in many industries, but are only on the verge of being fully utilized within biological production systems. The Broad Institute is pursuing these techniques to use them as a competitive advantage over other genome sequencing centers to meet the NHGRI’s ultimate vision, “to cut the cost of whole-genome sequencing to $1,000 or less, which will enable the sequencing of individual genomes as part of routine medical care.” Although many of these applications are new to the Institute, the employees and the laboratory environment, they prove to be extremely useful. Throughout the course of this internship, process re-engineering and six-sigma problem solving techniques have been utilized to improve processes and attempt to understand the variation that is inherent with any biological system.

The following section will discuss the use of these techniques, the novelty of them in this environment and the cultural change that is required to successfully implement processes that utilize lean production.

3.1 Preserving Quality in DNA Transformation

Sample misregistration problems fall directly on the workers within the MBPG. This is due to a number of reasons, including:

- Lack of data systems tracking samples.
- Numerous organism types in production.
- Operational system that lacks quality controls and error prevention.

The work in MBPG is manual, time consuming and can be very confusing. The processes have been developed and deployed to maintain the technological lead that the Institute has over competing genome sequencing centers, but the processes are not built for flexible operations. Processes lack segregation of different organisms and equipment that separates the work of different employees. In addition, multiple sample transfers occur without systems in place to prevent a worker from transferring a well to an incorrect plate.

Quality problems within the MBPG have a very significant impact on the organization downstream due to the amplification of material through the production process, as can be seen in Figure 6.
Given the relatively low cost of the process and very high cost associated with quality errors, it is critical that quality measures are put in place in the MBPG to prevent large scale quality excursions from the process downstream.

On June 19, 2005, a large-scale quality error occurred that was attributed to the MBPG. To assess the situation, the problem was approached with a structured problem solving analysis.

To begin solving the problem, a definition of the problem was created. The goal was to prevent sample misregistrations within MBPG. In order to do this, process maps were created for all MBPG process flows and areas of high risk were highlighted for further evaluation.

Figure 7 illustrates an overview of the process flow for the Broad Institute and MBPG within the operations. As mentioned in the introduction, the MBPG is broken into three areas; DNA preparation, DNA vector ligation, and transformation. Each process has specific roles, and dedicated technicians. It is rare that these technicians have the skills necessary to perform other jobs, and work is completely handed off between the areas.
After analysis of the different groups, it was highlighted that the riskiest process is the one associated with DNA transformation. In addition to being a highly manual and repetitive process, the transformers handle multiple organisms in a single day and use common equipment to process and store materials throughout processing.

The transformation process was chosen to be the focus of the improvement and was studied in much greater depth. One of the outcomes of the high risk process was that transformers had been blamed for several misregistrations in the past. These accusations were based primarily on a best guess, but the actual cause of the misregistration was unable to be determined due to the lack of data available.

3.2 Analysis and Improvement Methodologies

An important aspect of problem resolution is to understand both the problem and the tools available to solve the problem. Given the techniques that have been developed through Six Sigma, lean manufacturing, and other quality and problem solving techniques, there are a broad range of tools available. It is important to understand the capabilities and applications of the methods to make a decision that best suits the circumstances.

In the quality problems at the Broad Institute, it was most valuable to use the methods of process mapping, failure mode element analysis (FMEA), operational redesign, and
quality systems implementations. These methods will allow the users understand the process and the inherent risks associated with the current processing technology. With that understanding in place, the process can be designed to gain better outcomes and allow for prevention of future occurrences of the problem.

3.3 Transformation Process Mapping

Process mapping is used to visually understand what happens during a process. These maps are used to highlight processes, decision points, and inventory locations. As one analyzes a process in greater detail information can be gathered on what types of information is collected, how long it takes to complete and activity, and what types of materials are required for processing. These steps require numerous interviews with the employees and multiple observations of the process. Without taking the time to walk the process or do the processing for a period of time, it is unlikely that an accurate map of the process and decision points can be made.

During the weeks following the quality error, the transformation process was studied intently. All data, processes, equipment and transfers were recorded and placed in several visual formats. These maps were then used to understand the process and highlight important quality factors, such as; common equipment, common processing time, possible sample misregistration points, unclear processes, and processes that were conducted differently by different employees.

With a greater understanding the process, it is relatively easy to evaluate areas of wasted time and work as well as areas where the process is not robust. Figure 8 and Figure 9 represent how the transformation process was mapped. These process maps included the data that is recorded, as well as, a map of where they have to walk to complete each transformation.
3.4 Highlighting Risky Processes in Transformation

Once the transformation process was thoroughly understood, it was necessary to evaluate the Broad Institute’s greatest quality risks. In order to complete this part of the process, a team of transformers, management, operations specialists and members of the MBPG group was formed. The team spent four hours walking the process, from the maps that had been created previously. At each point the team brainstormed quality errors that could occur at each process step.

Once the point had been exhausted of ideas, the team continued to the next process step and repeated the steps. This was done until the process had been completed and the team had gone past the point of hand-off to core sequencing.

3.5 Re-engineering the Transformation Process

Process mapping yielded a wealth of information about the process, the steps associated with the process and where possibilities lie for sample misregistration opportunities. With that information, the process could then be designed for improved usability with fewer quality errors and less time each worker must spend manually recording data.

The approach used for the process redesign required all members of the transformation team to meet and talk about possible improvements in the working area and a discussion of the risks that the analysis had highlighted.
All FMEA risk points were discussed and the team brainstormed ideas about methods of avoiding these problems through operational changes, data gathering, or complete removal from the system. The data was then collected and prioritized through a risk analysis that associated risk levels (3 being highest and 1 being lowest) for possible misregistrations. A summary of the risk analysis is depicted in Table 3.14.

It should be noted that the data was taken in a group meeting. This data would have been more effective if it had been collected by each individual. This method would have allowed the project manager to define significance between risk factors of each issue.

<table>
<thead>
<tr>
<th>Case</th>
<th>Issue</th>
<th>Likelihood of causing 1 plate mis</th>
<th>Likelihood of causing multiple plate mis</th>
<th>Likelihood of occurrence</th>
<th>Total Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Labeling in difficult to read samples</td>
<td>Equipment</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Relies on transfer steps</td>
<td>Process</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Jackknife when brought into box</td>
<td>Risky</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Test tube was mishandled</td>
<td>Risky</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Glutinous information is not simultaneous</td>
<td>Process</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Common prices for Doing</td>
<td>Equipment</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>No back-up when lack of sample assumed</td>
<td>Equipment</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Outside the workflow</td>
<td>Process</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1 hour waiting required in incubator plates</td>
<td>Risky</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>Regular unexceptional processing error</td>
<td>Process</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>All back-up data are done for historical purposes</td>
<td>Risky</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>No back-up on serological database</td>
<td>Equipment</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>Other problems due to serological database failure</td>
<td>Process</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>Schedule and reporting the day and process is in place</td>
<td>Process</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>Work of mass schedule, not tube in inventory</td>
<td>Risky</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>Transfers put out by database mass schedule not production</td>
<td>People</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>Transfers are done of material with different staff</td>
<td>Risky</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>All transfers in the same inventory box</td>
<td>Risky</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>No back-up on serial plate</td>
<td>Risky</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>Transfers put out by database</td>
<td>Process</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>Master schedule is improper</td>
<td>Equipment</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>Replicate put out by error</td>
<td>People</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>Replicate was put out by database</td>
<td>Process</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>Replicate was put out by database</td>
<td>Process</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>Replicate was multi-problems</td>
<td>Process</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>26</td>
<td>Cumulative plan</td>
<td>People</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>Cumulative plan</td>
<td>People</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>No back-up on serial plate</td>
<td>Risky</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>No back-up on serial plate</td>
<td>Risky</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>30</td>
<td>No back-up on serial plate</td>
<td>Risky</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3. FMEA analysis of misregistration failure modes within the transformation process.

The primary problems associated with transformation processes were the following:
- All samples had hand-written labels that were difficult to read.
- Incubator processes and inventory locations had poor sample segregation.
- There was no information technology in place to track transformer processes.

To address the issues, operations improvements focused on putting short-term solutions in place in a timely fashion and planning a long-term solution that would gather process data and alert the transformers of problems with production.

**Improving sample labeling**

To address the problem of transformers not being able to read the labels on the tubes at the freezer inventory location, the Broad Institute purchased a software program that allowed the labels to be printed on a local printer and then placed on the top of the tubes.
This allowed the workers to see things more clearly, but still did not prevent the possibility of grabbing the wrong tube because, although improved, the organism labeling system has very similar labels for different organisms. An example of handwritten labeling can be seen in Figure 10.

![Figure 10. Handwritten tube labels.](image)

**Poor sample segregation**

As was highlighted in the FMEA, there were multiple points in the process where in-process inventory were placed in the same location. These points include;

- Freezer inventory locations
  - Unprocessed inventory
  - Excess, processed inventory
- Incubation steps
  - Post electroporation
  - Overnight growth plates
- Labeling areas
  - All plate labels and barcodes use the same printer

**Freezer inventory segregation**
To address this issue, all projects of different organisms were segregated and the inventory boxes were audited on a weekly basis. Prior to this change, all samples and organism types were kept in the same box. Also, projects that had been completed were not cleaned out on a regular basis. This allowed for a large misregistration opportunity with the possibility of eight or more organism types being represented in the same box.

New processes required sample segregating and dilution inventory maintenance on a regular basis. A new policy was implemented requiring that all leftover dilutions be discarded after transformation. The leftover inventory was not as necessary as it had been in the past because the entire DNA sequencing process has gained improvements in sequencing efficiency, which often leave leftover DNA dilutions after processing. This new policy prevented the re-entry of leftover inventory. The implementation of this policy reduces the possibility of misregistration by removing samples from inventory for processing and never replenishing them with excess material. The improvements can be seen in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>Before Freezer 5S</th>
<th>After Freezer 5S</th>
<th>% Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boxes in Freezer</td>
<td>369</td>
<td>102</td>
<td>72%</td>
</tr>
<tr>
<td>Tubes in Freezer</td>
<td>19188</td>
<td>3468</td>
<td>82%</td>
</tr>
<tr>
<td>% older than 6 months</td>
<td>72%</td>
<td>14%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 4. Freezer 5S improvements.

Improving inventory management at incubation steps

Within the MBPG, incubation steps are required to allow the cells to recover after the electroporation process which shocks the cells so they take on the desired DNA insert into the plasmid. In addition to recovery, incubation steps are used to optimize the speed of the amplification process to grow colonies of cells overnight.

The samples were initially placed in the same inventory locations for these processes due to the cost and size of current bench top incubators that are rated for molecular biology purposes and the amount of space required to hold up to 500 growth plates during times of high production volumes. The growth plates are placed overnight in an unorganized warm room that contained expired and unnecessary material. Figure 11 illustrates the organization of the room.
To improve upon these systems, the Broad Institute searched for incubators that were small enough that each transformer workstation could have one or two incubators on their lab bench for post-electroporation processes. Unfortunately, all laboratory equipment incubators were too large and costly for this possibility, so the search was broadened. Eventually, the laboratory purchased small incubators that are traditionally used for hatching eggs. These incubators were slightly modified and are now able to hold one transformation set. The incubators were tested to measure how well they held temperature and were allowed to be run with production. Each transformation workstation will now have two small incubators that segregate transformation sets. This eliminates sample misregistration risk during incubation. An additional benefit is gained because the samples never have to leave the transformers bench until they are ready for overnight processing in the warm room.

To address the overnight incubation sample segregation, the room was first cleaned. Some material in the room was dated more than six years and multiple hazardous materials such as broken glass, razor blades, and broken ceiling fixtures were removed from the room. After that, the room was segregated into several areas that allowed each transformer to have their own space for growth of plates overnight. Additional shelving options were researched, but none fit production needs.
Labeling areas

The current process requires that all labels be printed on the same printer. These labels are the start of the data tracking system for a library, and if these are placed on the wrong set of plates, all of the data will be swapped with another transformers work. To prevent any quality problems associated with this, the transformers are asked to print two sets of labels. The first set going on the plates during transformation and the second set going into their laboratory logbook.

Initially, the solution looked as though each transformer would have different labeling machines at their workstations and labels would be printed directly to the lab bench. This proved to be difficult for a number of reasons from an information technology and operations standpoint, so the solution was changed.

The new system, which will be focused on in greater detail in the information technology section uses the same labeling process, but alerts the transformer of a problem with the set and prevents them from moving forward with their work. This system will feed into the sample management network and prevent quality problems through user interface alert messages.

Information Technology requirements

As the Institute grew into a high-throughput DNA sequencing laboratory, the automation required information technology systems to be in place. These systems used barcodes to control sample management within the automated process and understood which DNA samples were being processed at each time. Due to the manual processes within the MBPG, these systems were not implemented in the group and are inherently the reason that sample misregistration problems often go undetected.

After the initial misregistration, it became obvious that the samples would need to be tracked. Due to the types of materials that the transformers use, normal barcodes are not feasible on tubes because the transformers must be able to see the ligation and the labels do not fit on the tubes.

In order to find a solution that would be suitable for the process, current technologies that are used by other life science and pharmaceutical companies were studies. The two technologies that were focused on were radio frequency identification (RFID) and two-dimensional barcodes.

The cost and technological barriers associated with RFID made it unattractive for current implementations, but is being investigated for other inventory management purposes at the Institute.

Two-dimensional barcodes have a number of high-throughput enabling technologies, including: accurate, handheld barcode readers, clear tubes with dimensions that were similar to current process tubes, and data that was very similar to other data entry points within the sequencing process.
The tubes have a slightly different form factor than those used previously, but have small barcode labels on the bottom of each tube in Datamatrix format. This labeling system allows each tube to have a unique identification that is tracked in the inventory management system. This unique identifier can be linked back to the original DNA sample that produced the material, therefore allowing the samples to be identified by the automation system.

There are a number of benefits to this type of technology and sample management system. First, all samples will be uniquely identified. This will allow the Institute to identify samples that are in the incorrect boxes and allow for future identification of poorly labeled tubes for auditing purposes. Second, data can be tracked to the original sample that produced the DNA used for processing. This allows the Institute to improve troubleshooting abilities for; bad ligations and problematic libraries. Finally, the technology has been developed and has many standards, allowing the institute to choose equipment from several vendors based on which one suits its needs most.

The sample tracking system is based on the premise that the ligation is entered into the automation system and associated with a uniquely identified tube. The material in the tube then goes through the processes of; dilution creation, electroporation, and sample plating before reaching the automated portion of the sequencing operations.

At each process step, the sample is transferred to a different container required for the process such as, electroporation cuvettes or agar plates. Each of the containers must now be uniquely identified and the transfers can be tracked through the use of information technology.

To uniquely identify a dilution, the Institute uses a new two-dimensional bar-coded tube that is identical to the ligation tube with the exception of the identifying barcode. Electroporation cuvettes must also have unique identification and the Institute chose a standard one-dimensional barcode placed on a non-electrode side of the cuvette. The agar plates had already been uniquely identified because they are the start of the information tracking system that has been used for many years at the Institute.

At the transfer steps, the transformer will scan the barcode of the source container, complete the transfer step, and then scan the barcode of the destination container. This transfer is recorded in a database and matches the new container to the sample history, therefore, identifying the specific organism and library. An example of the scanning process can be seen in Figure 12.
The system alerts the transformers of any problems associated with the containers that are being scanned. The problems highlighted that are most important for quality production are summarized in Table 5.

<table>
<thead>
<tr>
<th>Error</th>
<th>Resolution Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container is already associated with a library</td>
<td>Discard container and obtain a new one</td>
</tr>
<tr>
<td>The source container has already been transferred</td>
<td>Rescan and stop process if it errors again</td>
</tr>
<tr>
<td>Agar plate library not associated with cuvette library</td>
<td>A misregistration has occurred, do not proceed</td>
</tr>
<tr>
<td>The container has already been used</td>
<td>Discard container and obtain a new one</td>
</tr>
</tbody>
</table>

Table 5. Summary of system errors and required resolutions.

This system will give the transformers real-time feedback on their processes through the use of a tablet PC that is placed at each workstation. In addition to the wireless PC, barcode scanners for both two-dimensional barcodes and high-throughput one-dimensional barcodes are used to minimize time necessary to scan the containers.

As a final quality step, the agar plate unique identifier is already associated with a library and project. If the agar plate project history does not match the history from the tube, the system will not let the transformer continue with the process. This will highlight a misregistration near the source of the problem and allow for real-time problem resolution. This will prevent quality errors to be passed downstream in the process, to be sorted out later.

Additionally, the system had features built in to allow for real-time quality data for process control parameters. Because of manual systems that had been used by the transformers to track data in their logbooks, all electroporation and plating failures were recorded in the books, but not in a central location. In order to obtain the data, one needed to go through each of the transformers books day by day and record the information, making the assumption that all data was consistent, accurate, and recorded the same way across all people.

Quality control in an organization like the Molecular Biology Production Group is one of the most important parts of their production process. MBPG is near the start of the process and has little processing cost when compared to the other departments within the organization. Additionally, quality problems that leave the MBPG are not recognized until the DNA has been fully processed, requiring significant material cost and constraint processing time. The importance of improving quality within MBPG with regards to eliminating wasted processing costs and throughput time is illustrated in Table 6.

<table>
<thead>
<tr>
<th>Lost Consumable Spend if Quality Errors are Caught In-line</th>
<th>% of Process Cost</th>
<th>Lost Consumable Cost</th>
<th>Lost Processing Time at Constraint</th>
<th>Ability to Detect Mixups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caught After</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBPG</td>
<td>4.5%</td>
<td>4.5%</td>
<td>0%</td>
<td>Yes</td>
</tr>
<tr>
<td>Core Sequencing</td>
<td>92.7%</td>
<td>97.2%</td>
<td>0%</td>
<td>No</td>
</tr>
<tr>
<td>Detection</td>
<td>2.3%</td>
<td>100.0%</td>
<td>100%</td>
<td>No</td>
</tr>
<tr>
<td>Finishing</td>
<td>0.0%</td>
<td>100.0%</td>
<td>100%</td>
<td>Yes</td>
</tr>
</tbody>
</table>

32
Table 6. Cost of quality errors passed by MPBG.

To begin gathering this data, the capability was added to record the barcode of failed electroporations and failed plate growth. With the technology that has been added to the workstations, these data recordings can be accomplished real-time in a matter of seconds. The system can then provide reports that track the number of failures for plating and electroporation with a several failure mode choices. This data can then be characterized and the group can be alerted to trending quality problems and when quality is becoming out of control.

The capability provides the user with failure rates on agar plates each morning that is provided in a report to all operations managers. This capability allows them to track data based on which technicians and machines had processed each sample. This data is used to track process parameters and allow the MBPG to find problems and prevent them more quickly in the future.

During discussion in the structured problem solving portion of this thesis, an example of a large quality excursion will be discussed. Keep in mind that the data to do that analysis took more than fifteen hours to gather and analyze to find out if trending had occurred and the sources of the data were not known to be accurate. This will be a powerful tool as the MBPG moves further along the path of continuous improvement and process control.

Current processes at the Broad Institute require a great deal of manual data tracking and take up considerable amounts of time. Due to the increase in required productivity per transformer, the extra data tracking work and learning curve associated with a new system causes barriers for implementation, as will be discussed later. The system needed to be designed in such a way that the user could learn it quickly and have few errors when processing the material. It is necessary that the system take up very little space on the laboratory bench because many of the processes require large materials such as, ice buckets, agar plates and tip boxes. These requirements added difficulty to the design of the workstation.

3.6 Implementation Problems

3.6.1 Background on the culture

The process redesign had a significant change on the workflow for the transformers. The significance of the changes are illustrated in Table 7.
<table>
<thead>
<tr>
<th>Change Group</th>
<th>% of process changed</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT System Required Changes</td>
<td>100% change, no previous IT support</td>
<td>Addition of a tablet PC to the workstation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Addition of automated tracking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Addition of barcode scanners</td>
</tr>
<tr>
<td>Equipment changes</td>
<td>50% change. 5 out of 10 types of equipment changed.</td>
<td>2D barcoded tubes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Benchtop incubator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of different cuvettes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of different tips</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elimination of deep well plates</td>
</tr>
<tr>
<td>Process Changes</td>
<td>60% change. 6 out of 10 major processes</td>
<td>S.O.C. addition process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plating process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scanning transfers (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recording quality data</td>
</tr>
</tbody>
</table>

Table 7. Summary of the changes made to the transformation process.

In addition to the considerable changes that were going to take place for the transformers jobs, there were a number of significant cultural challenges that made for difficulty with implementation.

The Broad Institute has a unique culture that fosters innovation within molecular biology. This innovation resides with the scientists at the Institute who develop processes, automated machinery, IT infrastructure, and scientific breakthroughs. Innovation is the key to the Broad Institute’s success, as described by Rob Nicol, Director of Sequencing Operations, “I believe that our (Broad Institute’s) competitive advantage is our ability to innovate and quickly implement new, cutting edge technology.”

Although innovation is a key to the success and competitiveness of the Institute, this culture is not fostered throughout the organization. The innovation primarily takes place in the development groups and is quickly transferred to the laboratory. Laboratory workers are trained, but rarely have the ability to improve the processes after they have been successfully implemented.

The transformation process is a good example of this. The process was developed and passed to the operations group. Over the past few years, there have been very few changes that have taken place within the process and there has been little turnover from members of the group. In addition to the lack of turnover, the members do not tend to move to other areas of the process, such as DNA preparation and ligation. The skills required to do the transformations have been mastered by the current group of transformers, and as the transformation group is the bottleneck within MBPG (not the Broad Institute as a whole), they are also the workers who have the least amount of spare time for employee and skill development.

This has resulted in a group that produces a consistent product, but also does the same work each day and has done that for a number of years. Through this, the management team expects there to be no problems within the transformation process, because it is viewed as the process which the workers should have mastered and that there is little excuse for mistakes, as stated by a MBPG member, “The transformers are very
experienced and there is no excuse for them making these mistakes.” Thus, when there is a mistake, the transformers are quickly blamed for the problem that occurred and management and the rest of operations spend a great deal of time looking over their shoulders to make sure that they do not make more mistakes.

Over time, this has led the transformers to be very defensive against management, change, and even the abilities of the other transformers. One of them said this about their coworkers, “I would like to make sure that this system tells us who made the mistake because I know that the other transformers make all of the mistakes.”

These cultural barriers made gathering necessary information and making changes to the processes difficult because the transformers feared the following;

- The new process would take time to learn.
- The new process would cause them to make mistakes.
- The new process will highlight failures in their work.

These fears have also made the transformers be very timid about giving feedback on their processes. If they are trying to protect themselves against being attacked for something that they do by the management. Two of the transformers even have extra systems that they have used to keep track of information better than the current system requires. That highlights the fact that there are holes in the information gathering system, as well as the operations about how the process is run.

### 3.6.2 Redesign and Pilot Implementation

At the start of the redesign process, the manager of the transformation group highlighted one of the workers, Transformer A, to work on the process redesign. Transformer A was then asked to participate in the process and accepted.

After several meetings with all of the transformers and Transformer A in particular, it was difficult to get information about the groups preferences on the way that changes could be done. As part of the project, many options and pieces of equipment were proposed. These choices included;

- New incubators
- New tubes, tube caps, and tube holders
- Scanning tools
- Tablet PC preference
- User interface choices
- Redesign of the workstation layout

The majority of the decisions were avoided by the transformers and made by the project manager. These choices included the mounting for the tablet PC, barcode scanners, and incubators. At the time, this was not an issue, especially since the initial workstation would be a pilot and would be improved upon after the test.
Once the IT system was ready and the equipment had arrived, the pilot workstation was created. The initial pilot was scheduled to run for two weeks, requiring one transformation set each day for a total of ten transformations worth of data. The pilot would include changes that were made to the transformation process flow as well as recording data through the IT interface.

The Institute allowed Transformer A to spend half of the day on one set of transformations and complete the normal transformation process the rest of the day. The advantages of this method were twofold. First, the decrease in capacity would be reasonable to deal with and additional transformation help could pick up the extra one or two sets per day that would be required for production during pilot.

Second, as issues were raised about the pilot, especially with the IT system, changes could be made in the afternoon and loaded into the system overnight. This would allow the pilot to have consistent feedback and improvement of the system, such that it could be retested the following days.

In order to complete the pilot this way, there needed to be a commitment from the operations group and the IT department that they would provide the resources necessary to run the pilot with Transformer A and make the system changes required on a timely basis. These resources were allocated and the pilot was scheduled to begin September 25, 2005 and end on October 6, 2005.

In late September, the group stated that they would not be able to give the necessary resources to complete the pilot due to personnel issues. This required that the pilot be moved two weeks to begin on October 9, 2005.

On October 3, 2005, the MBPG began experiencing significant failure rates after the transformation process. The group began to troubleshoot the problems to find a root cause and found that the cells that were being supplied to the institute were having lower transformation efficiency and in many cases were yielding zero colonies. Electro competent cells are required to complete the transformation process and take several days to produce and ship to the Institute. The Institute worked closely with their supplier and began receiving new batches of cells, but the yields had been so low that the inventory at the operations bottleneck, the detectors, had gotten extremely low.

This required the management team to make the decision to put all resources toward transformation to keep the buffer from running to zero and the pilot was put on hold until further notice. In addition to the pilot being put on hold for reasons associated with the cell failures, the transformers were required to increase production significantly and the scrutiny of their work increased as well. This made the transformers even more defensive of their work, especially since they were now required to gather extra data on each transformation so the project team could find the root cause of the quality issue. As mentioned previously, detailed information about this issue and the resolution will be discussed in the Structured Problem Solving chapter of the thesis.
To meet business requirements, the pilot dates were changed to November 21, 2005 through November 23, 2005. These dates were chosen because the inventories upstream from the bottleneck had increased and staffing within the group was able to support the effort. However, the pilot was now only consisted of five transformation sets worth of data; one the first day and two sets each subsequent day. The IT department would not be able to support changes to the system during the pilot, so any errors that occurred would need to be dealt with during the pilot. Although not ideal, it was necessary for business purposes to make the changes.

On November 21, the pilot began in the morning with Transformer A completing the first set. It became obvious that there were several problems with the setup that was chosen. First, the amount of work required on the tablet PC made it necessary that the PC be in a ergonomic position. The initial arm used to mount the PC did not make this possible and would need to be changed. Second, there were a number of issues with the IT system that allowed the transformer to double enter codes or forget to start new transfer steps when it was necessary. Finally, the scanners that were chosen because they would be easy to use and had wireless capability were neither easy to use, nor capable to connecting wirelessly to most computers. Figure 13 illustrates the original transformation workstation and Figure 14 displays the initial pilot workstation.

![Figure 13. Transformation workstation prior to redesign.](image)
These problems made the pilot transformations take a great deal of time. Transformer A was frustrated, but had little constructive input to say about the process, often showing reluctance to the changes with comments, “You all will do what you want no matter what. I will have to stay here longer to spend all this time scanning. It will be terrible.”

After the pilot was completed, the data was analyzed and presented to Transformer A and is summarized in Table 8. The data showed a few small problems with the IT system that were fixed the following week, but primarily focused on improvements to the workflow with the redesign.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>n</th>
<th>Errors</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>1</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Dilution Step</td>
<td>7</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Cuvette Transfer</td>
<td>62</td>
<td>2</td>
<td>3.2%</td>
</tr>
<tr>
<td>Plate Transfer</td>
<td>120</td>
<td>6</td>
<td>5.0%</td>
</tr>
<tr>
<td>Cuvette Failure</td>
<td>2</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Total</td>
<td>192</td>
<td>8</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

All errors were attributable to IT problems and resolved.

Table 8. Transformation redesign pilot quality data.

Since the transformers were not used to changes in their processes and engineering redesign, the workers had never participated in a pilot. It was explained to the transformers that the pilot was the initial design for the new process, but would change after highlighting the inadequacies of the design. Those faults would then be changed...
and another pilot would be run if the changes were significant, until the final design took shape. These were new ideas to the transformers and they did not believe that this was the process that we were going to use because previous process changes were pushed on them by the development groups with little say in how it worked of impacted their work.

At the end of the meeting, Transformer A became very frustrated because she believed that the pilot system was going to be what was implemented as a final process change. The frustration of that prospect finally forced them to speak. For fifteen minutes, Transformer A gave feedback on every part of the process, the problems with it, the lack of need of certain aspects of it, the changes that would be required that were not taken into account and any other information that they could.

From that point on, Transformer A was very helpful in designing the next version of the workstation, depicted in Figure 15. The major changes that took place were:

- Changes to the IT format, including colors, pictures, and alarm boxes for problems.
- A uniquely designed, ergonomic swivel stand for the tablet PC.
- A wireless mouse to avoid needing to use the table PC pen.
- New scanners that are mounted and used for high throughput scanning.
- Changes to the electroporation cuvettes that are used in the process.

![Figure 15. Transformation workstation after 2nd redesign.](image)

### 3.7 Outcomes of Project Implementation

**Project Summary**
The goals of the project were to;
- Reduce sample misregistration in transformation processes to zero occurrences.
- Create a system that alerted the transformers to misregistration risks.
- Reduce the effort and time required to prevent errors and create high quality transformations.

The project resulted in the following;
- Reduction in sample misregistration occurrences.
- A user friendly IT system that gave instant feedback to the transformers about their work and any problems associated with it.
- A reduction in time required for lab notebook data tracking and overall transformation time.
- Significant improvement in operations workflow and reduction in possible misregistration opportunities.
- Data tracking for quality issues within MBPG.
- Roadmap for further implementation of IT to the beginning of the DNA preparation process.

3.7.1 Data Supporting Results Claims

Misregistration Data

In the seven months prior to implementation of the new process on a pilot workstation, there were a significant number of misregistration errors at the Broad Institute. The monthly data from Organism Checker has been compiled in Figure 16. Notice the downward trend toward a reduction in misregistration errors.
Of the errors that occurred, many number of them were believed to have happened prior to the automated processes based on information that was analyzed from the data logging machines. None of the misregistrations listed in that group had a full, root cause analysis done, but all were attributed to be the fault of the transformation process.

Post-implementation misregistration data has not been collected to find the significance of the improvement from transformation redesign. This data will not be available until after this document has been completed. Figure 17 displays the daily mismatch and low confidence sequence data. There is some a toward a reduction in large sample misregistrations after the process redesign and changes were put in place, but there has not been root cause analysis on the problems that have occurred, so a reduction due to the implementation of the redesign cannot be proved.
**Organism Checker Mismatch Data**

**Figure 17. Daily Organism Checker misregistration data, May 2005 – December 2005.**

*IT user interface provides instant feedback*

The IT system that was implemented has many improvements over the previous processes. Prior to the project implementation, there was no way for an outside source to tell the transformer that a mistake had occurred, that the ligation dilution was labeled incorrectly, or that the transformer was using the wrong set of labels.

The new system alerts the user to problems such as:
- Containers that have been used before
- Organism mismatch – preventing misregistration
- Transfer problems

The user also knows what organism they are working on because it is shown on the screen as they complete their work. The new system also allows a user to find out all the containers that the sample had been in previously, allowing for increased troubleshooting ability and possibly increased quality.

*Workflow improvements*

The redesign was successful in improving the workflow for the transformers. This allowed the majority of the work to be completed at the transformer workstation and
eliminated a significant portion of misregistration opportunities. Figures 18, 19, and 20 summarize the changes that took place during the redesign.

**WGS Transformation Redesigned**

- **Step 1:** Print out master schedule
- **Step 2:** Get plates from cold room
- **Step 3:** Put plates in the warm room
- **Step 4:** Sterilize desil and other tools
- **Step 5:** Get plates from current production in 20
- **Step 6:** Get dilution from current production in 20
- **Step 7:** Put cells in 400oucher
- **Step 8:** Put cells andelectroporation cures on ice
- **Step 9:** Bind dilutions in cures
- **Step 10:** Electroporate specimen
- **Step 11:** Incubate to 60 min
- **Step 12:** Incubate specimen (40 min)
- **Step 13:** Make labels at workstation
- **Step 14:** Compare labels with master schedule
- **Step 15:** Get plates from the warm room
- **Step 16:** Place plates in the warm room
- **Step 17:** Put plates in warm room again
- **Step 18:** Write down information and barcode lab book
- **Step 19:** Next day, when counts are estimated
- **Step 20:** Next day, white counts are estimated

**Figure 18.** New walking map with significant reduction in required processes and walking distance when compared to initial process illustrated in Figure 7.

**Figure 19.** Representations of the transformation process before (left) and after (right) redesign.

The workflow has been improved and there is a decrease in the amount of time that a transformer spends doing their work. This was done by the elimination of a number of process steps that are currently used. There is also an expected future benefit of manual
lab notebook data tracking elimination after the system has been implemented and tested across all transformer workstations.

**WGS Current Summary**
- 6 swap points
- 11 work locations
- 9 common use areas
- 8 sample location changes

**WGS Proposal Summary**
- 1 swap point (83% improvement)
- 5 work locations (55% improvement)
- 4 common use areas (55% improvement)
- 3 sample location changes (63% improvement)

**No Sample Re-entry Points**

Figure 20. Summary of improvement data from the transformation redesign.

*Quality Data Tracking*

Quality data has not been effectively tracked in the area and much of the data is not tracked at all. Due to the quality issues that happened during the implementation, it became obvious that this system would be even more useful if it created an interface that forced quality tracking.

The system supports the tracking of two types of data that were not tracked previously:
- Electroporation failures
- Agar plate failures
- Colony count (white counts) data

The benefit of tracking electroporation failures will allow the MBPG to better understand the cost of a failure and the percentage that it happens on a regular basis. This will allow the group to find out who is the best at electroporations and begin to train other members on best practices that the transformer has found. It has also been hypothesized that many transformations with low yields from agar plates also had a greater number of electroporation failures. These theories can be tested and a large amount of data can be collected to understand when processes are trending out of necessary quality ranges.

Agar plates are the final product of the MBPG and must maintain a specific volume to feed the inventory buffers before detection. Failure rates on agar plates are relatively low with sporadic periods of very high failure rates that can reach 75% for several days. These failures can be due to a number of causes, such as; low cell colony count, contamination, high cell colony count, and no growth. This data has not been tracked historically and is not available after the plates have been disposed of. The new system allows the user to choose a failure mode and scan all plates that were part of that failure type. This process can be completed in less than one minute each day and will allow for data to be tracked and easily accessible.

As mentioned prior, the new system gives the MBPG the capability to track their colony count data. This is important in understanding the control of their processes. In the past, the number of colonies on a plate was estimated by one of the operations managers.
Later in the process, the number of colonies that were found and picked by the automated picking machines were recorded. These numbers have not been tracked and compared historically due to the differences in information systems. The new interface allows the data to be compiled in one report and compared. Five days worth of data has been compiled in Table 9. It compares the number of plates, the estimated number of colonies, the number of found colonies, and the number of picked colonies. Control settings have are currently being developed for control charts.

<table>
<thead>
<tr>
<th>Date</th>
<th>Estimated Colonies</th>
<th>Colonies Found</th>
<th>Colonies Picked</th>
<th># Plates</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/24/2006</td>
<td>3695</td>
<td>8072</td>
<td>2457</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>1620</td>
<td>2625</td>
<td>829</td>
<td></td>
</tr>
<tr>
<td>3/26/2006</td>
<td>3185</td>
<td>6728</td>
<td>1986</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>738</td>
<td>1867</td>
<td>515</td>
<td></td>
</tr>
<tr>
<td>3/27/2006</td>
<td>2730</td>
<td>5020</td>
<td>1559</td>
<td>838</td>
</tr>
<tr>
<td></td>
<td>1384</td>
<td>2027</td>
<td>508</td>
<td></td>
</tr>
<tr>
<td>3/28/2006</td>
<td>2117</td>
<td>4235</td>
<td>1465</td>
<td>277</td>
</tr>
<tr>
<td></td>
<td>668</td>
<td>1686</td>
<td>488</td>
<td></td>
</tr>
<tr>
<td>3/29/2006</td>
<td>1703</td>
<td>3290</td>
<td>1218</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>761</td>
<td>1284</td>
<td>460</td>
<td></td>
</tr>
</tbody>
</table>

Table 9. Colony count data for five days from the MBPG.

Statistical Process Control charts are part of future implementation goals so that the group can be alerted to quality problems that the Institute is unaware of. This will be a valuable data resource when trying to find root cause of a problem because the start time of the quality issue can be tracked with high confidence data.

**Foundation for future tracking**

During the redesign and implementation process, it was important to choose vendors and systems that could be installed to other portions of the MBPG. For instance, the ligation process is very manual and requires flip-cap tubes to do their work. The types of two-dimensional bar-coded tubes that were purchased for the transformation process would not be suitable for ligation processes. Due to this, the Broad Institute chose a vendor that will develop the product for future sale to the Broad Institute and the market as a whole. This will allow for the IT systems to be brought back further and the ligation process to be able to implement the changes without a great deal of process change. After that, the IT system can be pushed upstream to the DNA preparation group.

In the redesign project, all of these were considered to allow for a smooth transition of other processes, allowing the Broad Institute to expand upon the data that it collects and improve the quality and problem resolution functions of the Institute.

**3.8 Conclusion**
Lean manufacturing techniques, specifically process redesign has made significant improvements across many industries. The techniques were applied to a genomics sequencing center and resulted in an improved process flow, process understanding and data collection.

The application of these techniques within an organization requires planning, teamwork, and a commitment to constantly improving and revising processes and ideas. Successful implementation of process redesigns can create significant improvements to cost and quality of a process.
4. Controlling Variability in the Molecular Biology Production Group

The operations at the Broad Institute are subject to significant variation. This variability is due to the complexity of the processes, the variation within the materials supply, the differences with how the workers complete their tasks, and the inherent variation within biological processes.

This chapter will focus on the variation that is caused by materials supply within the Broad Institute and define the impact that the variation within that supply can have on the quality of the processes. Additionally, the chapter will discuss the chosen method to minimize the variability within the MBPG and highlight a significant problem and resolution that was caused by materials supply variation.

4.1 Technical Background – Enzymes and Library Construction

The process of creating a cloned DNA library such that data can be produced from detectors is an intricate process that requires numerous reactions to occur with high efficiency. Additionally, the replication of the DNA relies on the ability of the competent *e. coli* cells to replicate the DNA, therefore requiring biologically healthy cells.

The process of taking sheared DNA fragments and creating libraries that can be replicated biologically by competent cells, called ligation, consists of several primary process steps. These process steps include several reactions and length separations to make sure libraries with the lowest insert size standard deviation. The main processes are; ethanol precipitation, linker ligation, and gel size fractionations. These processes are briefly described below:

- **Ethanol precipitation** – A process commonly used to concentrate DNA from aqueous solutions. It is the process of precipitating nucleic acid molecules by ethanol plus salt.
- **Linker ligation** – Linker ligation is the process in which the ends of the DNA fragments are standardized. This is caused by a reaction in which an enzyme repairs the ends of the DNA fragments and standardizes them for insert into a plasmid.
- **Electrophoresis gel fractionation** – This is the process by which the molecular biologist trims the undesirable strand lengths from the population. This is done by pulling the fragments through an agarose hydrogel by inducing an electric current. After seventeen hours, the fragments have sufficiently separated themselves and the molecular biologist can remove and purify the desired strand lengths.

The processes have been developed to use enzymes efficiently and maximize the number of quality reads at the detectors.

Despite the efficient process design, the process only provides a 5% yield for usable DNA after the ligation process. Much of this is due to loss within reactions, elimination due to undesirable strand lengths, or loss due to purification after gel electrophoresis. Table 10 quantifies process losses at each quality check.
Table 10. Loss data for the ligation process within MBPG.

To prevent further losses, it is important that reactions work efficiently and the cells are healthy. When looking at this in terms of the materials selected to create the reactions and replicate the plasmids, it is important to understand how the materials can be affected by the environment with which they are placed.

According to Sambrook, et al\textsuperscript{21}, to create a cDNA library, there are basic, necessary materials that are required. A list of them and their storage requirements, if any, are shown in Table 11.
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinomycin D</td>
<td>Store at -20C in foil wrapped tubes</td>
</tr>
<tr>
<td>S-Adenosyl-l-methionine</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>dCTP</td>
<td>Aqueous solution</td>
</tr>
<tr>
<td>ATP</td>
<td>Store at -70C in aliquots</td>
</tr>
<tr>
<td>Bacteriophage I arms</td>
<td>Store at -20 in aliquots</td>
</tr>
<tr>
<td>Bacteriophage T4 DNA ligase</td>
<td>Store at -20 in aliquots</td>
</tr>
<tr>
<td>Bacteriophage T4 DNA polymerase</td>
<td>Store at -20 in aliquots</td>
</tr>
<tr>
<td>Bacteriophage GT4 polynucleotide kinase</td>
<td>Make fresh just before use</td>
</tr>
<tr>
<td>beta-mercaptoethanol</td>
<td>Make fresh just before use</td>
</tr>
<tr>
<td>Deoxyribonucleoside triphosphates</td>
<td>Store at -70C in aliquots</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>E. coli</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>E. coli DNA ligase</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>E. coli DNA polymerase I</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>EcoRI</td>
<td></td>
</tr>
<tr>
<td>M. EcoRI methylase</td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td></td>
</tr>
<tr>
<td>IPTG</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>KCl</td>
<td></td>
</tr>
<tr>
<td>Linearized plasmid DNA</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>MgCl2</td>
<td></td>
</tr>
<tr>
<td>NaCl</td>
<td>Store at -70C in aliquots</td>
</tr>
<tr>
<td>(NH4)2SO4</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>NotI</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>Sal I</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>Oligo(dT)12-18</td>
<td>Store at -70C in aliquots</td>
</tr>
<tr>
<td>Packaging extracts for bacteriophage I</td>
<td></td>
</tr>
<tr>
<td>Phenol:chloroform</td>
<td>Store at 4C in dark bottles</td>
</tr>
<tr>
<td>Poly(A)+ RNA</td>
<td></td>
</tr>
<tr>
<td>Reverse transcriptase</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>RNAase H</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>RNAase inhibitor</td>
<td>Store at -20C in aliquots</td>
</tr>
<tr>
<td>Sephadex G-50</td>
<td>Room Temperature</td>
</tr>
<tr>
<td>Sepharose CL-4B</td>
<td></td>
</tr>
<tr>
<td>Sodium acetate</td>
<td></td>
</tr>
<tr>
<td>Synthetic linkers</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td></td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td></td>
</tr>
<tr>
<td>Tris base</td>
<td></td>
</tr>
<tr>
<td>Tris CL</td>
<td></td>
</tr>
<tr>
<td>X-gal</td>
<td>Store at -20C in tubes</td>
</tr>
</tbody>
</table>

Table 11. Critical materials and storage requirements to create a cDNA library.

One can notice that there are a number of enzymes that are used in the process in addition to competent cells. There are other materials that are temperature sensitive as well and require storage at extremely low temperatures.
To understand why the storage requirements have been chosen, it is important to understand the role of the material and how they are impacted by temperature. Since the ligation process requires enzymes, the focus of the next section will be on the role of enzymes in the reaction and how they can be impacted by environmental conditions.

Enzymes are the biological catalysts, usually proteins, that help carry out nearly all chemical reactions in living systems. These catalysts control reactions within organisms that allow them to function. Like other catalysts, the role of an enzyme is to decrease the activation energy required for a reaction to proceed. According to thermodynamics, reactions will proceed if there is a total decrease in the free energy that happens after a reaction. Unfortunately, there is often a barrier to starting the reaction, which is called the activation energy. Enzymes decrease the thermodynamic barriers required for a reaction to take place and bring it within the energy ranges that are feasible for living organisms. Figure 21 provides an illustration about how enzymes decrease the energy necessary to begin a reaction.

![Energy Diagram]

Figure 21. An illustration of how an enzyme reduces the energy required to start a reaction.

Once the activation energy is low enough for the reaction to take place, the reaction begins to happen at a rate that is governed by kinetics. Reaction rates decrease over time due to; substrate depletion, approach to equilibrium, enzyme activation, and other artificial causes.

Enzymes can be optimized by their environment to create the highest efficiency. The methods of optimizing enzymes that are often used are those that are relative to the buffer solution that the enzyme is contained in as well as external factors such as temperature at which the reaction takes place. For high-throughput DNA sequencing processes, many of these factors are already taken into account as corporations that supply the enzymes manufacture them at the optimal pH and buffer content. Within the protocols, the
external environment is controlled through incubation processes and thermal cycling equipment that is used in the laboratory.

This optimization should result in a high enzyme efficiency, which will produce a reaction that allows the enzyme to perform its function. These reactions are necessary to create strands of DNA that have specific end sequences such that they can be placed in a vector to create a plasmid.

As one can imagine, it is important that the enzymes are high quality and high efficiency. If they are not, a reaction may not fully take place. The results of a low activation enzyme could be any of the following:
- Reaction does not occur.
- Only part of the reaction occurs.
- Library is contaminated and must be recreated.

Since the DNA yield between reactions is very low with efficient reactions, it is important to make certain that these reactions are not hampered by enzymes with low activity. When this occurs, the DNA may not be trimmed to the desired length and a significant portion of the target DNA could be lost.

This would mean that the researcher would lose significant portions of valuable DNA due to a non-reactive enzyme. This increases the price and time required for researchers and lab technicians to perform their experiments.

To understand what causes an enzyme to become less active, it is important to understand what an enzyme is. Most enzymes are proteins and they catalyze reactions. These enzymes have an affinity, or non-affinity to certain parts of a DNA strand. As DNA themselves, the enzymes will become active given certain conditions and perform an action on the DNA. If an enzyme goes through a conformational change and becomes denatured or deactivates, the enzyme is not able to perform this activity.

Denaturation is usually classified in accordance with external condition inducing it, for example, by heating, by cooling, by pressure, by acids or alkali, or by denaturants such as urea or guanidinium salts. These all can cause a conformational change that will denature the enzyme and prevent it from performing the duty that it is supposed to.

Denaturation of an enzyme is often a reversible process, but it requires specific conditions that are not always met in a laboratory environment, especially one focused on production. If an enzyme has low activity due to denaturation, the Broad Institute would recreate the reaction and throw away the enzymes that are suspect to be bad.

Given the importance of the reactions to take place at the Broad Institute and the purchasing policies for the Institute; such as purchasing enzymes and competent cells from a supplier that specializes in them, we can assume the optimum storage pH and sterilization of the enzyme is high.
It is important to understand how temperature can impact the activity of an enzyme. As per Privalov, it can be shown that enzymes are rarely in an intermediate state. This means that once the conditions for denaturation occur, they will occur quickly to those impacted and over a short range of the conditions. Research has shown that denatured enzymes lose their catalytic function, and therefore do not provide an environment for reactions to proceed.

These criteria relay the importance that materials should be kept at -20°C for storage and materials such as, Bacteriophage T4 DNA ligase and RNAase inhibitors should not be used if they have gone through several freeze-thaw cycles.

In addition to the loss of activity if an enzyme reaches a thermal barrier that causes denaturation, enzymes also have decreases in activity related to freeze-thaw cycles. As stated for many enzymes by several researchers, enzymes lose activity through freeze-thaw processes. The amount of activity lost relies heavily on the buffer solution in which the enzyme is stored and the type of enzyme that it is. Some enzymes lose significant portions of their activity from freeze-thaw cycles while others seem to be robust. This phenomenon is illustrated in the following diagram. In Figure 22, the y-axis is representative of the activity of an protein and the x-axis represents the number of times the protein has been through freeze-thaw cycles.

**Figure 22. An illustration of enzyme activity loss after being subjected to numerous freeze-thaw cycles prior to use in a reaction**

These data show that enzymes can lose activity through external temperatures being placed upon them and freeze-thaw cycles. To better understand how this can impact Broad Institute operations, it is important to understand how a decrease in activity can cause negative results for the Institute. The following highlight some outcomes associated with low activity enzymes:

- An enzyme loses the ability to catalyze a reaction and therefore the reaction does not occur. This may result in the scrap of one or several libraries, resulting in significant worker-hour loss.
An enzyme may only partially catalyze a reaction. The results could be the following:

- It may result in the scrap of one or several libraries, resulting in significant worker-hour loss.
- The worker may choose to complete the library or libraries, but may be subject to significant loss of DNA yield at the end of the process.

Considering that there are numerous reactions that take place throughout the process, this loss can occur in several portions of the process flow. Thus, the variability in the quality of enzymes significantly increases the risk of a non-reaction and the likelihood of a low-yielding library. Low yielding libraries can result in the following:

- Low transformation yield. This is caused by a low concentration of plasmid DNA during electroporation. This can be caused by a low DNA concentration in the ligation solution.
- Recreation of libraries. If DNA yield is low, the Institute may require that multiple libraries be created to satisfy project needs. Each library accounts for 40 worker-hours.
- Increase in transformations required. If DNA yield is low and ligations cannot be diluted significantly, the transformers may be required to transform many more plates, requiring a significant risk to sample misregistration and manual work-hours.

To minimize this variability, the Broad Institute purchases the majority of critical materials from professional suppliers. The suppliers have expertise at creating the materials and have developed a strong relationship with the Institute. Additionally, the materials group within the operations organization has specific protocols and quality control tests before materials are delivered to the laboratories to be used in sequencing operations. Although these processes have been optimized to some degree, there is still significant variation due to handling of the materials by both the materials and laboratory groups.

Competent cells are also critical for the DNA sequencing process. The *E. coli* cells are specially treated to be used during the electroporation process where they take in the new plasmid DNA and replicate it numerous times over the following seventeen hours. Like the enzymes, these materials are supplied from an outside vendor in sterile packaging conditions and created in a highly controlled environment. Once received, the cells are stored in -80°C freezers until they are used in production.

Cells are stored at this temperature to maintain quality of the material. As noted in several laboratory manuals and articles, the cells can be used for up to six months with little loss in efficiency if they are stored below -70°C. Additionally, Mazur and others have shown that freeze-thaw cycles have a damaging effect on cells and can injure them. The cryogenic process is very stressful on the cells because of the rapid thermal changes and creates other stresses induced by ice crystal formation in the solution that holds the cells.
It has been shown that freeze-thaw cycles damage the cells and decrease the efficiency at which the cells can transform the ligations. Baum\textsuperscript{35} notes that the success of a transformation in the electroporation process relies heavily on the fitness of the cells at the time.

Academic research suggests that freeze-thaw cycles can have a negative impact on the fitness of the cells\textsuperscript{41,42,43}. This is done by reducing the efficiency at which cells produce transformations as well as the yield of living cells by the time they reach transformation. In addition, since the success of a transformation relies heavily on the fitness of the cells, it can be induced that freeze-thaw cycles and long-term storage can have a significant impact on the efficiency and quality of transformations.

The Broad Institute uses very aggressive electroporation parameters for cell transformation. According Sambrook, the optimal transformation efficiencies are produced when the death rate of the cells are roughly 50-75%. Since the Institute is focused on high throughput of DNA sequencing, they target aggressively to produce optimal transformations. If the cells have been injured throughout the storage process, there are several things factors that can have a negative impact on production. Some of these factors are:

- Cells have already been too injured to withstand the electroporation stresses and die with a spark discharge.
- A significant number of cells have already been severely injured or have died and another 50-75% mortality rate decreases the cell population too much to produce significant transformation efficiency.
- Cell membranes do not adequately recover from the electroporation process and the cells aggregate preventing an efficient transformation.

All of the possibilities listed above have a direct impact on the labor required to produce the transformation agar plates for the automated portions of the process and greatly decrease the control that the operators have over the process. Additionally, a great deal of the variation that is introduced to the electroporation process comes from the storage, delivery, and age of the cells that are used.

As will be described in more detail in a later example, the Institute production operations had a significant problem that was caused by the injured cells. As the team worked to find the root cause of the problem and take corrective action, it was determined that the cause of the cell problem was that the packaging used for the cells did not provide enough insulation to prevent the cells from going through freeze-thaw cycles during delivery periods. These cycles impacted the health of the cells and decreased the transformation efficiency of production significantly.

4.2 Preventing Material Degradation

The information on the previous pages has described the drivers for products of the processes that take place in the MBPG. For instance, the product of each enzyme reaction and gel electrophoresis step is the amount of DNA, or yield that can be obtained
for the next step. It has been shown that the quality of the enzymes have a direct impact on the product of a reaction and are a critical material for production. It has also been shown that enzymes and other materials within the process, including competent cells can be significantly affected by the conditions imposed on them during storage and delivery.

To minimize the variability and maximize the quality of the reactions, it is important to provide critical materials that are of high quality and have been exposed to minimal temperature changes throughout the storage and delivery processes. As will be shown in the following section, variation within the process was significant and materials had the capability of going through several freeze-thaw cycles before they were consumed during production. Solutions have been implemented to minimize this problem for nearly all materials delivered to the laboratory.

Addressing the Issue

The Molecular Biology Production Group has a great deal of inherent variability in their processes. This inherent variability is caused by their use of living organisms to produce their products. It would be valuable to understand this variability such that biological processes can be better understood and so measurements can be obtained that to minimize variation outside of the expected, inherent process noise.

Part of this thesis is to reduce the variability within the group such that the Broad Institute could move toward a better understanding of their process capabilities and eliminate quality problems as they begin to occur. Since there is currently very little data that is collected in the group, it has been difficult to produce a baseline of variation, but work done previously by Matt Vokoun, LFM '05 standardized processes and eliminated the use of different equipment\(^4\). This thesis focuses on finding areas of high variability and attempting to minimize the variation within those areas.

4.3 Understanding the Existence of Variation

In order to reduce variability, one must show the existence of variability. Within the MBPG there are many things that can be measured. These metrics include; production volumes, yields, operating costs, and some quality data. Also, if one takes all of the data from laboratory notebooks, one can track the variability of parts of the processes. Some of the variation within the group can be seen in Figure 23\(^4\).
Figure 23. Ligation production cell colony count data mean and standard deviation.

Since the data shows that there is variation, it is important to understand the causes of the variation. To understand the sources of variation within the process, a cause and effect diagram was created. Each of the causes were ranked in a failure mode element analysis risk matrix to prioritize projects. Each of the projects have possible tools listed that can help improve or eliminate the variation. These diagrams are in Figures 24 and Table 12 respectively.

Figure 24. Cause and Effect Diagram of process variability.
<table>
<thead>
<tr>
<th>Cause</th>
<th>Cause large variation</th>
<th>Cause variation to many processes</th>
<th>Cause variation to many people</th>
<th>Likelihood of occurrence</th>
<th>Total Risk</th>
<th>Method to Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadequate quality tracking</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Quality IT</td>
</tr>
<tr>
<td>High supplied material variation</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Kanban, QC</td>
</tr>
<tr>
<td>Using different materials for same processes</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>5S</td>
</tr>
<tr>
<td>Wrong or inadequate supplies</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>5S</td>
</tr>
<tr>
<td>Expired material usage</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Kanban, inventory</td>
</tr>
<tr>
<td>No problem resolution and prevention</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>54</td>
<td>Six Sigma</td>
</tr>
<tr>
<td>No root cause analysis</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>54</td>
<td>Six Sigma</td>
</tr>
<tr>
<td>Biological Variability</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>54</td>
<td>Inherent</td>
</tr>
<tr>
<td>No IT assistance for operations</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>54</td>
<td>IT</td>
</tr>
<tr>
<td>High interuser tool variability</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>36</td>
<td>5S, preventative maintenance</td>
</tr>
<tr>
<td>Old, poor equipment</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>36</td>
<td>5S</td>
</tr>
<tr>
<td>Dirty / contaminated process equipment</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>36</td>
<td>5S</td>
</tr>
<tr>
<td>Inconsistent training</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>24</td>
<td>Training protocols</td>
</tr>
<tr>
<td>No equipment and process qualification method</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>24</td>
<td>Standardized process change procedure</td>
</tr>
<tr>
<td>Inconsistent equipment</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>24</td>
<td>5S</td>
</tr>
<tr>
<td>No equipment monitoring systems</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>18</td>
<td>Quality systems, maintenance</td>
</tr>
<tr>
<td>Inconsistent protocols</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>16</td>
<td>5S</td>
</tr>
<tr>
<td>Little preventative maintenance</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>16</td>
<td>Maintenance</td>
</tr>
<tr>
<td>Neglect to detail</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>5S</td>
</tr>
<tr>
<td>High worker turnover</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>Training protocols</td>
</tr>
<tr>
<td>Poorly written protocols</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>Protocol improvements</td>
</tr>
<tr>
<td>Not following protocol</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>Training protocols</td>
</tr>
<tr>
<td>Wrong background</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>Hiring</td>
</tr>
</tbody>
</table>

Table 12. FMEA risk analysis for process variability.
Table 12 highlights a great deal of variability and what can be done to improve the situation. As one looks at the ranked risk factors and the tools that can be used to improve them, one can see why the prior project went through a 5S implementation. Of the twenty-three risk factors, seven of them can be improved through 5S, two of which are in the first four risk factors. After that, inventory/Kanban improvements, quality tracking and Six Sigma problem solving techniques become very important and useful tools to reduce the variation within MBPG.

Since variation exists and high risk areas have been identified, problems can be addressed objectively. This section will focus on incoming material variability.

### 4.4 Reducing Variation in Materials Supply

Once the variability in materials and reagents supply was chosen as a major contributor to variation within the MBPG, it was important to choose processes that reduce the variation of the supply. Figures 25 and Table 13 are in depth analysis of what could cause material variation. The analysis highlighted several contributors to the variation in the process that had not already been addressed through prior improvements.

![Cause and Effect Diagram](image)

**Figure 25.** Cause and Effect Diagram of material supply variability.
<table>
<thead>
<tr>
<th>Cause</th>
<th>Risk to 1 plate</th>
<th>Risk to many plates</th>
<th>Risk to many ligations</th>
<th>Likelihood of occurrence</th>
<th>Total Risk</th>
<th>Method to Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using bad material</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Kanban</td>
</tr>
<tr>
<td>Allowing significant temperature changes to the materials</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Freezer Kanban</td>
</tr>
<tr>
<td>Poor delivery equipment</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>81</td>
<td>Freezer Kanban</td>
</tr>
<tr>
<td>Material delivery taking too long</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>54</td>
<td>Kanban</td>
</tr>
<tr>
<td>Poor vendor processes, resulting in poor quality</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>36</td>
<td>Quality control</td>
</tr>
<tr>
<td>Broken, inaccurate freezers</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>36</td>
<td>5S</td>
</tr>
<tr>
<td>Wrong materials are used</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>27</td>
<td>5S</td>
</tr>
<tr>
<td>High age and age variability of materials</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>24</td>
<td>Kanban</td>
</tr>
<tr>
<td>Leaving freezer open too long</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>18</td>
<td>5S, Freezer Kanban</td>
</tr>
<tr>
<td>Poor inventory replenishment process</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td>Kanban</td>
</tr>
<tr>
<td>Dirty warm and cold storage rooms</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>18</td>
<td>5S</td>
</tr>
<tr>
<td>Not using first-in, first-out inventory</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>Kanban</td>
</tr>
<tr>
<td>High inventory levels</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>Kanban</td>
</tr>
<tr>
<td>Expired material is delivered</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>Material supply IT</td>
</tr>
<tr>
<td>Material is delivered with wrong ID (supply mixup)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>5S</td>
</tr>
<tr>
<td>High material variability from vendor</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>Quality control</td>
</tr>
<tr>
<td>Using own materials vs. laboratory supplies</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>5S</td>
</tr>
<tr>
<td>Inability to return bad material</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>5S</td>
</tr>
<tr>
<td>Bad Incubators</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>5S</td>
</tr>
<tr>
<td>Broken autoclaves, allowing contamination</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>5S, preventative maintenance</td>
</tr>
<tr>
<td>No IT assistance preventing material supply errors</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>SAP or other material management system</td>
</tr>
<tr>
<td>Reusing leftover material</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>5S</td>
</tr>
<tr>
<td>Poor quality of internally produced materials</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>5S, quality control</td>
</tr>
</tbody>
</table>

Table 13. FMEA analysis of materials supply variability risk.
These areas included;
- Expired material usage
- Thawing of frozen materials – especially enzymes
- Handling / contamination of materials
- Variability in age of materials / reagents used
- Freezer temperature variability

After highlighting these as possible issues, they were addressed individually.

**Expired Material Usage**

Initial inventory was taken of all supplied reagents. The technician used the data to form diagram of the inventory box with the dates that the material was created. The age of the material is associated with a color, ranging from green (most recent) to red (expired). The age of the material is broken down into different color segments and the red indicates that the material is expired.

As can be seen from the two examples below, Figure 26 and Figure 27, the materials were generally organized well, but the age of the materials became significant. For materials with a shorter shelf life (such as 6XLD30), many of the tubes had already expired from use, as illustrated in Figure 23. This data highlighted that nearly 30% of the supplied materials in the laboratory had expired and that many of them were supplied in quantities greater than necessary for inventory purposes.

![Figure 26. Diagram representing the age of reagents in the MBPG. Dates represent the date that the material was made.](image)

![Figure 27. Diagram representing a second material with significant expired material.](image)
Variability in the age of reagents

Variability in the materials supplied ultimately leads to greater variability within the MBPG. After investigating delivery processes with the materials group, it became clear that the materials group does not supply expired materials because the inventory management system that is used does not allow a transaction of an expired chemical. This meant that the materials that were expired in the laboratory, had expired while on the laboratory shelf, not from being supplied bad material.

The variation in the age of the materials in the MBPG is significant. This is due to the fact that there is not a forced first in, first out system. Therefore, those materials that sit in portions of the box that people tend to grab first will be used first and some other materials may never be used. This causes a large variation in the age of materials in an inventory location.

A solution needed to be created that took into account inventory sizes that were consistent between materials based on time to use. It was also necessary to create a new system that required that the inventory be used in a first in, first out system so the standard deviation of time before the materials are used is consistent.

Thawing of frozen reagents and enzymes

Many reagents that are used in molecular biology are kept in frozen states at about -20 C. The reason for this is that the enzymes activate and have less consistent reactions when they are not stored at a very low temperature. To ensure quality processing and quality materials, thawing of reagents must be prevented. The current processes for materials delivery take a long time. In addition to the amount of time it takes to deliver the materials, many of the frozen storage reagents thaw on the cart because of the time necessary for delivery. In addition to the long time required for delivery, the volume of material supplied, often 25 – 50 μl aliquots, is very small and thaws quickly at room temperature.

Figure 28 illustrates the temperature changes of a sample during the materials delivery process. This was done by making an aliquot of a reagent, poking a hole in the top of the cap and sticking a thermocouple through it. The data was then collected every twenty seconds for the simulated time for delivery. The best case and worst case scenario data was collected for the time from when a reagent is removed from the freezer until it is replenished in the laboratory freezer. Notice that the materials thaw and nearly reach a steady state at room temperature.
It is necessary to create a system that addresses the issue of temperature control during the Broad Institute’s internal delivery cycle.

Handling and contamination of materials

For reasons stated earlier, it is necessary to minimize contamination of materials and reagents. Because the Institute does research on DNA, any organism, including the workers in the laboratory and airborne particles can contaminate a material. In addition, the contamination may cause the materials to activate, which increases the variability of the process.

Previous material storage bins had a great variation in the time that a material spent in the bin, but also had no way of protecting the materials from the surrounding environment. The aliquots were placed in bins that either had no lids, or the lids were never used because they were difficult to open and close. Also, there was little space between tubes in the previous inventory bins, so it was difficult to grab an aliquot without touching other aliquots.

It was important to find a system that made it easy to remove aliquots from the inventory bins and make it such that no material was left without a lid for more than two weeks.
Freezer Temperature Variability

Not only are temperature variability issues present with transportation of frozen reagents, but there are also problems with the variation of temperatures within a freezer. Each time that a door is opened, a great deal of heat escapes and the materials inside the freezer are exposed to these changes.

Due to the variability of freezer temperatures from being open and closed, it is believed that any materials in a commonly used freezer, including finished ligations that are older than six months should not be used because of temperature cycles. Materials and ligations that are critical to a project are stored in freezers that are opened rarely and only by authorized personnel with a key. This prevents archived material from being compromised from temperature cycles.

This concern brought additional requirements that the frozen materials be insulated from changes in the environment around them.

4.4.1 Design and Implementation of a two bin Kanban

Based on the analysis of the laboratory and the processes within the Institute, it became necessary to design an inventory management system with the following attributes;

- Small, consistent, usage based inventories.
- Ability to keep frozen reagents frozen.
- Easy to handle cases.
- Insulation for frozen reagents.

It was decided that the best way to do this would be to create a two bin Kanban system. Each bin would have a small inventory of no greater than two weeks. Doing this would ensure that all reagents are used, on average, within four weeks of when they are delivered. To determine bin size, all reagent usage data for the previous year was analyzed. The maximum, minimum, mean, median, and standard deviation of weekly usage were analyzed and the bin sizes were determined.

Using the criteria, the bin sizes are listed in the following table versus their original bin sizes. Table 14 displays a selection of materials with original material age, standard deviation, and weeks of inventory compared with the new Kanban system. As can be seen, the improvements were significant, including more than a twelve week average reduction in reagent inventory.

<table>
<thead>
<tr>
<th>Material Age (days)</th>
<th>Averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Material Age</td>
<td>83.7</td>
</tr>
<tr>
<td>Material Age St. Dev. (days)</td>
<td>26.1</td>
</tr>
<tr>
<td>Mean / Shelf Life</td>
<td>46.5%</td>
</tr>
<tr>
<td>Weeks Inventory</td>
<td>15.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Old Inventory</th>
<th>New Inventory</th>
<th>Weeks Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>16</td>
<td>5.0</td>
</tr>
<tr>
<td>52</td>
<td>28</td>
<td>3.9</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>3.7</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>4.4</td>
</tr>
<tr>
<td>25</td>
<td>40</td>
<td>4.3</td>
</tr>
<tr>
<td>25</td>
<td>40</td>
<td>4.2</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>4.1</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.8</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.7</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.6</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.5</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.2</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.1</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>3.0</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>2.9</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>2.8</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>2.7</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>2.6</td>
</tr>
<tr>
<td>24</td>
<td>16</td>
<td>2.5</td>
</tr>
</tbody>
</table>
Table 14. Reagent age improvements after the Kanban system was implemented.

The bins were chosen to be easy to use and provide space for the worker to easily grab what was necessary. It is important to note that the bins actually increased some of the inventories and take up more surface area than the previous system had. Although this sounds negative, the increase in surface area made it much easier for the workers to obtain the materials. It also allowed for enough space to close the lids on materials that were not in use. Figure 29 displays both the original inventory bins and the new Kanban bins.

Frozen materials used a similar two bin system with the same type of analysis. The difference was that frozen transport boxes were purchased so that the materials would be kept at -20 C. These boxes would be kept in the freezer at all times, except when materials were being replenished in them, or a worker was removing an aliquot for production use. The thermal boxes provide a vast improvement over the previous temperature changes that the aliquots were exposed to. One can see from the data below that the boxes do not allow the chemicals to thaw during the previous time required for delivery. In addition, the time required for delivery was cut by nearly 75% after the implementation of the Kanban system. Data is presented in Table 15.

<table>
<thead>
<tr>
<th>Average Delivery Time</th>
<th>Average Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Kanban</td>
<td>1:05</td>
</tr>
<tr>
<td>After Kanban</td>
<td>0:17</td>
</tr>
<tr>
<td>Improvement</td>
<td>72.7%</td>
</tr>
</tbody>
</table>

Table 15. Kanban delivery time improvement data (minutes).

For enzymes, which are even more critical to keep at a temperature, an extra level of insulation was used when they were being replenished. These carriers are kept in a -20 C freezer at all times except when in use. The temperature changes of an enzyme in a transport box is almost zero as can be seen in Table 16.
Summary of Kanban Temperature Change Data

<table>
<thead>
<tr>
<th>Degrees / minute</th>
<th>No insulation</th>
<th>Thermal Boxes</th>
<th>Thermal + Insulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>2.40</td>
<td>0.50</td>
<td>0.03</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>4.66</td>
<td>1.52</td>
<td>1.96</td>
</tr>
<tr>
<td>Minutes to thaw</td>
<td>3:20</td>
<td>49:40</td>
<td>&gt;&gt;60:00</td>
</tr>
</tbody>
</table>

Table 16. Reagent temperature increases during delivery processes using three different delivery methods.

The Kanban system was set up such that all empty bins were put in the same area that was outside of the laboratory. This allowed the materials group to quickly access the materials at any time without having to enter the laboratory, count materials, or open freezers. This point was also where the replenished materials were stored until someone from the laboratory placed them back into their minimart location.

Figure 30. Laboratory layout with respect to Minimart and Kanban locations.
Due to this, the materials and boxes could sit for a significant period of time at the replenishment point and would require that a freezer be placed there for the frozen reagents. To prevent the freezer door from being opened when it was not necessary, an Andon light was placed above the minimart. As can be seen in Figure 31, the replenishment point was organized, compact and color coded. The red area and red light signals to the materials group that there is at least one bin that needs to be replenished. The yellow light signals to the laboratory that at least one bin is waiting to be restocked in the laboratory minimart. If both lights are on, there are bins that require replenishment and require restocking. No lights indicate that there is no action that needs to be taken.

Due to the new bin inventory size, the stocks only need to be replenished once or twice each week and the materials delivery personnel began planning to pick the bins up when they were heading to do other tasks in the materials room. This allowed them more efficiently use their time and it decreased the time necessary for materials replenishment greatly.

Although it has been difficult to find data showing that this reduced the variability within the overall process of MBPG, the data shows that the variability within materials delivery and use has been reduced dramatically. As the group begins to implement systems that track quality data more accurately, it may be easier to characterize improvements and consistency of work within the group. In order to accomplish these improvements and understand the variability within the process, the following opportunities are available for variability reduction and data tracking.

![Figure 31. Kanban replenishment center.](image-url)
4.5 Future Areas of Focus to Reduce Variation

Person to person variability

The MBPG has a great deal of uncharacterized, interpersonal variability. This is due, in part to the training and inconsistency in the way that people do their jobs. Unfortunately, this variation has not been characterized to date and the data that has been gathered does not allow for quantitative analysis and trending of production from different people.

As an example, the transformers produce agar plates that have a specific number of colonies on them. The colonies are very time consuming to count individually and there is not automation in place to count the colonies. Due to manual time constraints, the data is estimated by a person glancing at the plate and giving a ballpark estimate. Generally, the data resides in multiples of 500. For instance, most counts are 1500, 2000, or 2500. In addition to being an estimate, the estimate is based on a transformation set. For instance, rather than having data that shows 14 of the 24 plates in a transformation set had colony counts of 1500, while 5 were 2500, and 5 were 500, the data will show that the colony count for the set (average of all plates) was 1500.

Gathering this data, although difficult will allow the Institute to better understand the inherent variability in their processes. As a way to obtain some quantitative data, the picking machines list how many colonies they picked from each agar plate and the data is collected by the tool automation. This may be able to provide an initial way of understanding the variation within the transformer daily output. It should be noted that the data obtained from the server is the number of colonies picked, which is less than the number of good colonies, which is less than the number of total colonies. Although this data does not correlate exactly with output and plate performance, it may be able to provide insight into the variation within transformation output.

The Broad Institute recently began tracking this data to better understand the variability within their process. The data that was described above can be seen graphically in Figure 32. Understanding this data will help the Broad Institute understand the variation that is caused by processes like colony count estimates.
Within the ligation area, many of the processes, including gel creation and gel cutting are done manually and are likely to vary greatly between individuals. Although this is the assumption, little data has been collected on this and the only available data is obtained by pulling it manually out of laboratory notebooks, which were 61% completed, as shown in Table 17.


<table>
<thead>
<tr>
<th></th>
<th>Ligator 1</th>
<th>Ligator 2</th>
<th>Ligator 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Input Points / Library</strong></td>
<td>140</td>
<td>140</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td><strong>Avg. Points / library</strong></td>
<td>107.1</td>
<td>77.7</td>
<td>77.3</td>
<td>85.5</td>
</tr>
<tr>
<td><strong>% Complete</strong></td>
<td>76.5%</td>
<td>55.5%</td>
<td>55.2%</td>
<td>61.1%</td>
</tr>
<tr>
<td><strong>St. Deviation Points / library</strong></td>
<td>16.9</td>
<td>24.6</td>
<td>13.4</td>
<td>24.6</td>
</tr>
</tbody>
</table>

**Table 17. Ligator lab notebook process data.**

Systems that can be put in place to record and trend the percentage of DNA recovered from each gel that was processed, as well as data that displays the total output of the completed ligation would be valuable. The Broad Institute is moving toward an e-labnotebook\textsuperscript{51} that will begin to record some of this information, but it has not been used by the laboratory workers yet.
As automation within MBPG continues and more workers have computer access at their workstations, this data will become easier to collect and display for the workers, allowing them to understand the variation within their processes and between people. Once this is done, workers can begin to understand the variation between them.

Standardizing materials and processes

The Broad Institute makes a great deal of their reagents and materials. This is done by the materials group in several laboratories within the building. These reagents are produced and checked for quality purposes, then supplied to the laboratory. These processes are an example of the vertical integration within the Broad Institute show the dynamics of the operations environment.

Although the Institute has produced these materials for several years, the mission of the operations organization is to produce DNA sequencing data at high quality and low cost. Since many of the materials produced within the laboratory do not have proprietary technology and are commodities within several major chemical producers, the Broad Institute could obtain more consistent materials that are produced in cleaner environments, specifically created for that type of production.

A cost and quality analysis should take place to understand the best use of resources for the Broad Institute and implementation of this could improve the quality of the sequence data produced.

Many processes within the MBPG are done manually. Examples of these processes are producing and cutting agarose gels. The production of gels takes place at the time of use and the gel area takes up a great deal of space in the laboratory. These items are available for purchase from large biological supply companies and would very likely produce more consistent results than production by workers in the laboratory.

Once the gels have been processed, the DNA must be removed from the gel. This is done by manually cutting out portions of the gel using razor blades. Attempts have been made by the group to create tools that make cutting the gels easier to do. Despite this improvement, there is a great deal of opportunity for the workers of the MBPG to innovate this process and make something that is easier to complete and more consistent.

There are numerous other examples of these causes of variation within sequencing operations, many of which could be easily understood and improved. As with any improvement in structured problem solving, the difficulty lies in finding a way to measure the baseline and improvement rather than creating a new process. Data collection and statistical analysis are important to all improvement processes and the implementation of changes.

4.6 Conclusion
Reduction of variability in materials supply can improve the quality of a process. These improvements can have an even larger impact if the incoming material depends on the environment that the material is exposed to.

Tailoring solutions, such as Kanbans, to one’s specific application can decrease the variability in the time a material spends on a shelf as well as the exposure of the material to external factors that hinder material performance. Tracking improvements in the outcome of reduced incoming material variability can be difficult and requires data prior to and after implementation.
5 Structured Problem Solving

This chapter will investigate a specific example of structured problem solving techniques that were used at the Broad Institute. The methodology, the steps, and the outcomes will be discussed to illustrate the improvements that can be obtained by using data to lead an organization to problem resolution.

This section is intended to provide a framework for using the method for future problem resolution to the reader.

5.1 Prior Outcomes of Quality Issues

The Broad Institute has made great strides in the field of molecular biology and high volume DNA sequencing. Throughout this period of time, the Genome Sequencing Operations have encountered numerous problems and worked their way through them. Despite this, the workers within the Institute rarely find a resolution to problems and fix the root cause of the issue.

This is due to a number of reasons, which primarily revolve around the culture of the Institute. As mentioned before, the culture at the Broad Institute is strong and the workers have great pride in what they produce. The operations are vertically integrated and the organization has a strong hierarchy. The setup of the organization has led to a culture that places all current initiatives on a high priority list that is constantly changing.

High priorities cause constant changes in focus of the workers and rarely allow time necessary to finalize current projects before starting the next high priority item. Also, within MBPG, data is difficult to collect with the current technical infrastructure and the necessity to have manual execution of procedures for DNA preparation, vector ligation, and transformation. The lack of central data locations make it difficult for workers and managers to have fast access to the data for problem solving purposes.

As with most dynamic environments, the failures that happen within the process often stop occurring within a few days. At that point, the team has just begun collecting and analyzing data and then the problem ceases to persist. This is a sign to the team that the problem has been resolved and was likely due to a number of causes that are attributed to issues including;

- Bad ligation
- Bad cells
- Hot or cold weather
- Someone in the group was sick
- That is the variability of biological processes

Although some of these causes may be valid, it is unlikely that the same causes reoccur in a controlled environment such as the Broad Institute. The workers then move to the next high-priority item and leave the problem unresolved without preventative measures in place to keep it from reoccurring.
In an effort to understand their processes better and create preventative measures, the Institute has invested in Six Sigma training for a number of the employees. This training is focused on teaching the structured problems solving DMAIC (Define, Measure, Analyze, Improve, and Control) process. Additionally, the MBPG has hired a new worker that will focus on process creation, process improvement, and problem solving abilities. The worker has a strong background in molecular biology and has the ability to use a different toolset that is necessary to solve some of the issues.

This new person, Researcher A, became involved in a quality issue that occurred soon after she began working at the Institute. Researcher A’s strong scientific background, the length and severity of the problem, and Six Sigma training provided an ideal opportunity to showcase the capabilities of the workers using the DMAIC process.

5.2 Quality Problems in Molecular Biology

Throughout 2005, there were several large quality excursions that occurred within the MBPG. The problems would cease to persist and the analysis of the issue would be halted. In early October, 2005, the MBPG began seeing a drop in the number of colonies that were produced in the transformation process. The cell colony counts dropped below 500 colonies, which is the failure threshold. For a period of two weeks, the group saw production failures ranging between 30% and 70%. This had an impact on downstream operations as there was a fear that the constraint of the Institute would run out of inventory, which would result in lost total throughput. Figure 33 is a graphical representation of plate failure rates within the MBPG.

![Graph of plate failure rates within the MBPG](image-url)
Figure 33. Failure rates in MBPG, June ’05 – October ‘05

To solve the problem, the group began theorizing and collecting data to try and explain the issue. These explanations often caused the management team to make changes in the processes to try and explain the reasons for failure. These changes included; changing processing parameters, using different, unqualified cell types, and changing ligations. None of these changes worked and the team was no closer to the root cause of the issue.

Led by Researcher A, the group defined the problem as low colony count plate failures. The group then produced a Cause and Effect Diagram, Figure 34, showing all possible causes of the failure with regards to processes, people, equipment, and materials.

Figure 34. Cause and Effect Diagram for electro-competent cell quality issues.

After completing the Cause and Effect analysis, the group was attempting figure out what the best way to attack the problem was and which to test first. The group began to order the problems based on intuition that supported their theories about the root cause of the problem. This produced a task list that was great and had little coherence across the members of the team. After a short period of deliberation, it was suggested that rather than listing them based on their opinions, they should focus on the likelihood of them being the problem.

To order the problems, they ranked the scores by the perceived likelihood of the failure mode impacting on plate or set and another ranking with the perceived likelihood of the failure mode impacting a large number of plates or sets. Another column was ranked with the likelihood of occurrence. All rankings were based on a scale of 1 to 3, 3 being the highest risk. These scores were multiplied to give a ranking number. Now, all of the
items were ranked with a score between 1 (1x1x1) and 27 (3x3x3). The risk analysis is illustrated in Table 18.

<table>
<thead>
<tr>
<th>Cause</th>
<th>Test</th>
<th>Risk to a single plate</th>
<th>Risk to all production</th>
<th>Likelihood of occurrence</th>
<th>Risk Factor</th>
<th>Owner</th>
<th>Estimated Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cells are past or nearing their expiration date</td>
<td>Check logbooks and remove expired cells</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>27</td>
<td>Mechele</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Cells have been stored in our freezer, and how long in materials freezer</td>
<td>Begin collecting data and compare received to use dates</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>18</td>
<td>Mechele</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Lab freezers don’t hold temperature well</td>
<td>Monitor freezer for 1 week once we receive probe</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>18</td>
<td>Kim</td>
<td>10/21/2005</td>
</tr>
<tr>
<td>Check-in time of lot takes too long</td>
<td>Ask Peter / Abdul and set restrictions, redesign process</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>18</td>
<td>Dave - must be in dry ice</td>
<td>10/21/2005</td>
</tr>
<tr>
<td>Electroporators are not pulsing correctly</td>
<td>Check voltage and resistance with outside meter</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>Jason - find out how to test</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Low DNA in ligation</td>
<td>PICO or pUC19 curve</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>Kim</td>
<td>10/14/2005</td>
</tr>
<tr>
<td>Materials freezer does not hold temperature</td>
<td>Monitor freezer for 1 week once we receive probe</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>Kim - give to Peter or Abdul</td>
<td>10/21/2005</td>
</tr>
<tr>
<td>Incoming cells are bad</td>
<td>QC testing</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>Mechele</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Electroporation settings are not correct</td>
<td>Compare data that we have, check to make sure machines have correct settings</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Dave - data</td>
<td>10/11/2005</td>
</tr>
<tr>
<td>Thaw time of cells is not correct</td>
<td>Ask transformers</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Cells are not put on ice immediately after removal from freezer</td>
<td>Ask transformers</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Lab freezer doors open too often</td>
<td>Freezer monitors and checklist</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Delivery process is too long or risky</td>
<td>Understand process, set requirements, and change process to meet the requirements</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Dave</td>
<td>10/21/2005</td>
</tr>
<tr>
<td>People return thawed cells to the freezer</td>
<td>Ask transformers</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>Mechele</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Materials freezer door opens too often</td>
<td>Freezer monitor and checksheet</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>Dave - talk to Matsi</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Outgrowth time does not meet specification</td>
<td>Ask transformers and look in logbook</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Warm room temperature is not correct</td>
<td>Use a thermometer and check against strip reader</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>Incubator temperature is not accurate</td>
<td>Check with a thermometer</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Dave</td>
<td>10/7/2005</td>
</tr>
<tr>
<td>S.O.C. is low quality</td>
<td>Check for use of different lots and QC by spreading on a plate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Kim</td>
<td>10/11/2005</td>
</tr>
<tr>
<td>Production volume becomes overwhelming</td>
<td>Check production volumes vs. problems</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>Mechele</td>
<td>10/11/2005</td>
</tr>
<tr>
<td>High salt concentration in ligation</td>
<td>??</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Kim</td>
<td>10/14/2005</td>
</tr>
</tbody>
</table>

Table 18. Electro competent cell failure mode element analysis (FMEA).

This list ranked the failure modes in relative importance and gave the group a better process for attempting to eliminate failure causes. Each failure mode, descending in risk level, was assigned an owner, specific actions that were necessary, and a date that the actions had to be completed.
As the group began to eliminate possible failure modes, they also gained a better idea of the data that was necessary to be collected and refined the data parameters. This data was collected by the team members, management and the transformers for a period of several days.

After obtaining the data, the FMEA had been reduced to a few possible outcomes. Further data analysis and presentation of the data in a good visual format pointed to the cause of the problem being the electro competent cells that were supplied by the Broad Institute’s main supplier. Due to the importance of the issue, a meeting was set up with the supplier and representatives came to the Institute to discuss the issue.

To start the meetings, Researcher A provided the suppliers with the analysis, and showed data that supported her claims. Rather than providing theories about the possible causes of the problem, the meeting began directly at the point of fixing the issue. The suppliers were impressed and convinced the analysis and were ready to work together with the Institute to solve the problem.

In addition to avoiding conflict and wasting important time, the supplier used the format that had been created with the fishbone and FMEA and produced the same information about their processes on the evening following the initial meeting.

The next day was spent planning timelines and setting up communication strategies so that the problem could be solved as quickly as possible. The FMEA analysis that the supplier created was the roadmap that they used to solve the issues.

Within a few weeks, the root cause of the problem had been found to be the packaging operations and specific requirements of the Broad Institute. These requirements were relinquished and the MBPG now receives the electro competent cells from the same packaging and process flow that all other customers receive the cells.

5.3 Outcomes from Structured Problem Solving

The most important outcome of using the DMAIC process was that the group was able to understand the root cause of the problem and put preventative measures in place to prevent a reoccurrence of the issue. This was a great example of how, despite the variability in biological processes, the cause of a problem can have something to do with shipping, handling, or packaging. The problem also showed the team members how to work through the process and saw the noticeable outcomes.

In addition to the final root cause, the group gained a better understanding of the electro competent cell delivery process and found problems with the temperature variability within the laboratory storage freezers. Temperature collection monitors have been placed in the freezers to monitor temperature and allow access to the data on a real-time basis to quickly test possible freezer temperature failure modes. Other equipment and processes have also been characterized and monitored due to this issue.
The employees of the Broad Institute also began working closely with the supplier to solve the problem. Once the problem was solved, several of the team members visited the supplier facility to learn more about the operations and discuss future business opportunities. The meetings resulted in more business for the supplier and a strong, open relationship with the MBPG. This will allow for future problems to be solved quickly and also provides the supplier and the MBPG with a better understanding of the needs and abilities of both organizations.

In summary, the structured problem solving process gave workers an opportunity to use problem solving tools effectively and work toward an important outcome. It created better monitoring and data collection techniques by highlighting processes and equipment that characteristically have little data tracking and the group built a better relationship with their supplier.

Since that problem, Researcher A and several other members of the group have used the DMAIC process to analyze other problems. Although many of the members are still learning about the tools available to them and the order in which to use them, they are beginning to understand problems in their processes more effectively.

5.4 Conclusion

The DMAIC process and other problem solving techniques can reduce the time that it takes to find a root cause to a problem and create a corrective action plan. An example of a successful outcome within an organization that is beginning to adopt structured problems solving methodologies has been described, including the final results. One can use the framework of the chapter as a format to begin applying the techniques to problems encountered within other organizations and laboratories.
6 Organizational Structure and Employee Development

This chapter will discuss the organization and culture of the Broad Institute and analyze the impact that it has on the ability of the operations organization to react quickly to new processes. In addition, it will provide recommendations that are tailored to the Broad Institute and the culture that it has developed through its leadership in genomic sequencing.

Background

The Broad Institute is a convergence of government, industry and academia. This has brought together some of the most talented researchers in a number of disciplines with a goal of eliminating and preventing human disease. The government has provided the funding necessary to promote programs that have little return on investment for public companies and huge barriers to entry. Employees with experience in industry have given the Institute the necessary resources to effectively use the machinery and efficiently sequence genomic data.

As mentioned prior, the organization has a very hierarchical structure that is clearly exhibited in Figure 35, the Broad Institute’s organizational chart.

![Figure 35. Genome Sequencing Organizational Chart.](image)

Most major initiatives and decisions are brought down from the hierarchy and projected to the sequencing operations groups. These initiatives sometimes have input from the workers doing the processing, but primarily they are set initiatives that must be achieved.
6.1 The Broad Institute – A Growing Organization

The Genome Sequencing Platform is at a critical turning point. The organization is continuing to grow and is implementing processes that are much more complex than those that were run previously. The complexity has to do both with the processes themselves and the number of different processes that are run at any point in time. The new complexity, the growth in employees, and the new problems that the Institute is seeing have required them to implement complex systems and processes to accomplish tasks.

These processes are new to the workers within the Institute because previously, work was accomplished through a series of informal discussions and changes that did not require documentation or communication of the change. The growing scale of the platform has required communication in the recent past, but still allowed many changes to take place based on informal communications and actions.

There have been growing pains associated with these changes. Primarily, the workers within the organization have been subject to more meetings than they had in the past. The employee view of the meetings is that they are often impractical and a waste of time. Additionally, the culture has shown resistance to attending meetings by not showing up on time or prepared for them. A student intern that was working at the Institute set up several meetings to get their project kicked off. The first two were cancelled because nobody attended the meeting or viewed it as important. After speaking with the internship sponsor, it was recommended that meetings be set up through the sponsor’s assistant so people knew it was a meeting of importance.

The requirement to set up the meeting through the sponsor’s assistant sent a key message to the invitees that it was an important meeting and they should attend. It also supports the claim of the strong hierarchy within the organization and the top-down approach to most initiatives. The middle management tends to carry out the orders of the upper management without question, which leads to a long list of high priority items for the employees on the floor and the direct managers.

As the organization continues to grow and implement structure, it will also be important to implement behavioral norms of the organization. This would be a document of expectations of the people working within the organization. These norms should be published and withheld so that all employees understand the expectations placed upon them as their jobs become more structured. Common norms of other learning organizations include: being on time and prepared for meetings, producing meeting minutes, adjourning meetings at scheduled times, blocking out lunch meetings or meetings on Friday evenings. Although not all of these norms apply to the Broad Institute, implementing norms tailored to the workers should help with the inevitable future of structure that will be required within the sequencing platform.

Another important aspect of the development of the organization will be to include the workers in the decision making processes and allow them to have an impact on which initiatives are chosen for the organization and how they will be implemented. Currently,
workers have little incentive or desire to make changes to their jobs because they have had negative feedback when they have done so in the past. This prevents the organization from developing the employees and reaching the potential that the workforce talent pool could achieve.

Without involving the workers on the floor, future initiatives, especially with regards to lean manufacturing and operations excellence will be difficult to successfully implement.

6.2 Six Sigma Training

A major initiative that is taking place at the Broad Institute is to train a number of the employees in Six Sigma processes and structured problem solving. As mentioned in a prior section, the results of a single project example are positive and show that the employees have the ability to implement and use the processes available to them. It can also be seen that the successful completion of training has been generally good at the employee level, as can be seen in Table 19.

<table>
<thead>
<tr>
<th>Number of Participants</th>
<th>Number Complete</th>
<th>Number on Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>39</td>
<td>68</td>
</tr>
<tr>
<td>Percentage</td>
<td>54%</td>
<td>94%</td>
</tr>
</tbody>
</table>

Table 19. Employee Six Sigma training progress.

This is a long term investment in the future capabilities of the Institute and employee development. It shows a commitment from management to improve the skills of the people at the Broad Institute and therefore, improve the overall capabilities of the facility.

The initiatives are brought down from the management and rarely come from the workers on the floor. In addition to that, the initiatives created at the top, are passed directly to the employees and have little buy-in from the middle management, although they typically support the initiatives.

Management support of programs dwindle after the initial implementation, but the employees feel that they have to complete it or else they will have a negative performance review as suggested by one worker, “It seems like we get in trouble for not doing things, but are rarely rewarded for doing something great. We just want to get the Six Sigma training done to avoid a bad review.”

This action has built a barrier between the employees and the management that should be reduced as the organization moves toward creating a workforce driven system of operations that can be flexible to customer demands make necessary changes without requiring management approval.

Giving the opportunities to all employees and encouraging them to make their own development plan will send a message to the employees that they control their careers. Many employees will not choose to go through the path and stay in their current roles, but more recent hires and high performing individuals may be attracted by the
opportunity to be considered for a different role after a reasonable period of time in their job.

These changes may also increase the satisfaction of the employees. Currently, the MBPG has a very high job turnover rate. This is detrimental to the group for a number of reasons, but primarily because the job procedures take a significant period of time to learn and the learning curve is steep. As turnover increases, the capabilities of the group and the investment in employee development programs yields lower returns.

6.3 Implementation of Structured Processes

As the Broad Institute moves toward a more structured environment, it must consider several important factors within all changes that are made. These factors include:
- Too much structure is unattractive to the current employee base.
- Create methods for all levels of employees to have input.
- Have some of the changes initiated from the ground level.

If the management team can tailor their efforts to the culture, they will be able to leverage the creativity and natural talents of their workforce. They will also get support from the workers by encouraging them to take a part in changing the culture. This will allow new structure, such as behavioral norms, to be taken on by the workforce as a whole. If the people are just told that they need to be on-time to meetings, but management does not role model this or allow the workforce to have influence, the culture will continue to perpetuate tardiness.

In addition to making sure that everyone is involved, the initiatives should take place from people within the organization. A number of changes within the sequencing operations platform have taken place based on internship projects focused on operational improvements. Many of these projects, including this thesis, have been successfully implemented, producing favorable results. The projects would be more effective if they were led by someone within the organization. These people have been referred to as an outsider-insider, or a person who understands the culture within the organization, yet can bring in new perspectives and creativity. An outsider-insider will be able to provide both the mindset that is necessary to see the changes that have to be made and will have the credibility and the network to make sure that the changes are implemented. Additionally, the person is assumed to have a long term investment in the organization versus that of a six month intern.

These changes will allow the Broad Institute to grow their employee capabilities, employee satisfaction, and the effectiveness of the Genome Sequencing Operations platform.

6.4 Conclusion

It is necessary for the operations organization to improve processes that encourage the flow of information as the headcount of the Institute increases. These changes will create
the structure necessary to transfer data, best practices, and knowledge between the employees. Examples of the processes that should be put in place are cultural norms, standards forms, and online data sources.

These processes will allow the workers at the Institute to use their talents efficiently and adjust quickly to changes in a medium sized organization. These changes will be difficult because of resistance that will be met by the culture that has been developing for several years. To be successful, management will have to commit to long-term support of the changes and have them driven by non-management members of the organization.
7. Sustaining Improvements

As mentioned previously, the Broad Institute has hired interns from a graduate program that focuses on operations and supply-chain logistics. These interns addressed operational problems within the organization. While the internships have been successful from a project standpoint, the primary investment in the interns is to expose the workers to people who have not been trained in biological processes, but have a better understanding of lean manufacturing and operations technology. The Institute is attempting to invest in their capabilities through the internships, as stated by Robert Nicol, “I would like to think that 80% of my investments are long term investments. I am less concerned with the outcome of an internship project and more concerned with the interaction that the workers here have with the interns.”

These interns have been placed into groups and work with the employees directly. Some years, the Institute has funded several internships allowing for collaboration and interaction between the interns and to the employees. These have been successful in the implementation of systems, but have had difficulty changing the culture of the organization to appreciate the improvements and changes that were made.

As an intern, it is difficult to find the incentives necessary to get the Broad Institute employees appreciate the value of many of the programs. Much of this is due to a lack of familiarity with other workers roles, but some of it can be attributed to differing values. For instance, an intern who is focused on operations may value the following attributes of an organization:

- Reproducible results
- Efficiency / organization
- Standardization of processes
- Low cost production

Although these may not be the values of every member of the program, many of the interns are likely to have this mentality. On the other hand, the employees of the Institute are scientists and have been trained in laboratories prior to their current jobs. The following are typically valued by a worker within the MBPG:

- Flexible processes
- Minimal value to the cost of production
- Minimal data tracking
- Personal organization that is not standard

The differences in values that the intern has versus the organization in which he or she is placed can make it difficult to implement change that will stay in the organization. In order to do so, incentives have to be created that relate to the workers in the group, but may not be inherently obvious to the intern. In addition to understanding the right incentives, it is important for the intern to create incentives for the employees to begin valuing some of the standardization and cost processes in a way that is acceptable within the culture.
In order to accomplish this, it is important for the intern to get an outsider-insider that believes in the work they are doing, supports the changes, and will work with them to make it amiable to the culture. Some of the internships in the past were more successful at creating this environment to pull change than others, but some current initiatives will be discussed in the following sections.

7.1 Sustaining 5S

The MBPG had an intern in 2004 that focused on organizing the laboratory and implementing 5S. During the internship, he changed the laboratory layout, made operational flow changes, labeled, and color coded the entire work area. To make sure that this was working and to track the progress, he created 5S check sheets that the employees were to use as an auditing process for their areas on a monthly basis.

After the intern left, the 5S check sheets stopped being used and some of the organization that had been implemented began to deteriorate. Many of the organizational metric gains that had been initially attributed to 5S also began to leave. This is shown in Figure 36.

![5S Checklist Scores](image)

**Figure 36. 5S checklist data.**

Due to the cultural shift that was required to implement 5S and make it part of what the group accomplished on a regular basis, the internship did not last long enough to fully influence the group. The responsibilities of 5S were left up to one of the managers, rather than a member of the group. This made it difficult because the commitments and goals that the manager has do not align with the goals of the laboratory with regard to 5S.
management goals are to make sure that resources have been allocated correctly, that people understand that tactic and strategic nature of their work and to communicate Broad Institute initiatives.

The workers have responsibilities of working in the laboratory and completing the work on a daily basis. The organization of the laboratory has a significant impact on the time that it takes to find materials and complete the molecular biology processes that they utilize. Since their work focuses on the using the materials and the equipment in the laboratory, they should be in charge of the 5S.

To maintain and create a culture around 5S, there is a need for a long-term investment in the processes. To improve upon the gains previously implemented by another intern, the current intern needed to show the laboratory the importance of 5S and monthly evaluation. As the internship ended, the responsibilities were handed off to the worker in the laboratory that was most proactive about fixing the problems that had been highlighted during the 5S audits.

This person also took the responsibility to reassign roles for the smaller, weekly audits of each workstation and followed up with the workers to make sure that these things were complete. Although the culture had not fully shifted to support 5S, significant ground was made by having one person take control of the 5S work.

The 5S checklist was also adapted to current processes due to inherent changes in the lab. These changes took place because many of the auditors that that some of the ratings were redundant or not relevant. Another group was in the process of implementing 5S and there was a need to have consistency across the organization for the monthly audit checklist. Several of the audit points were changed and the checklist was used for both MBPG and Core Sequencing.

Changing the culture of an organization, even if it is small, is often more difficult than most people anticipate. This difference in expectations of the implementation person and the group that the changes are focused on make it difficult to accomplish major changes, even in a reasonable amount of time and few technical difficulties standing in the way of the change.

The experience that has been gained from working with the MBPG and implementing changes that have worked well with those that have worked poorly can be summarized in a few points;
- Always listen and ask for input.
- Relate to people in a way that is important to them and aligns with their incentives.
- Be humble and never claim to understand something that you do not.

Following these simple points makes it easier to build relationships and understand the incentives that drive individuals. Once their needs are understood, it is much easier to tailor solutions that work well for them and fit into their daily routines. It also provides and environment where the workers feel free to change the system and improve upon it.
Collaboration and growth rarely happen if people are just trying to accomplish a personal goal, regardless of the goals of the group that they are working with. This prevents new ideas and systems from lasting within a group and changing how people do their jobs. Most importantly, it allows many people to close their minds to new ideas that may be very important for the success of their organization in the future.

Below are some quotes of workers at the Broad Institute relating to the internships of current and past students.

"It is very frustrating to be in the office with him because all he cares about is the thesis. It is obvious that he doesn't want to be here."

"I don't think that he understands that we are different. He believes that there is only one solution, and you have to do it that way."

"I think he did a good job, but I wish he were open to more ideas, or explained what he knows in a more comprehensive manner. This is all new to us."

As with many things, these are only symptoms of how people feel, but could be interpreted to show distrust in some of the internships and what they are attempting to accomplish. This distrust prevents organizational change by not influencing people to buy-in to the changes that are taking place. It should be noted, that not all internships were like this. Matt Vokoun, specifically, was spoken very highly of within the organization.

7.2 Extending Sample Tracking

The value of the sample tracking implementation within MBPG has been shown with marked improvements based on data shown previously. This data supports the fact that there was a need for a new system of tracking samples and that the new system has a vast improvement over previous systems. It also can be seen from the Figure 37 that the new system is not complete.
In addition to highlighting that DNA preparation and vector ligation do not have systems in place that track tubes and allow for workers to have the opportunity for sample swaps. This is also important to understand because the entry point into the informatics system is the last step in the ligation process. If this step has been done incorrectly, all of the tubes that come from that ligation will be misregistered.

Given this risk, an important part of the project was to implement a system that could be expanded upstream to include all processes after the DNA is received on site. In order to do this, the choice of the systems was important so it could allow consistency within MBPG. This goal was highlighted to the information technology specialists that designed the system that was used for sample tracking such that they could build expandability into their systems.

Most importantly, it was imperative to involve members of the group and especially the person who would take on the leadership role for the next steps of the implementation. This person could be involved in choosing the vendors such that they understood future needs of their products and good decisions could be made according to the strategic goals, not just the current implementation. The person needed to be involved in the decision making process and vendor choice process. The person should understand the service levels of the vendors that had been chosen.
All of these decisions could have been made by the intern alone, but would not have allowed for the organization to feel as involved in the process or know where to begin once more implementation needed to take place.

Rather than highlight a specific person, due to the changing needs of the organization, it was decided that all of the project goals and decisions would be communicated to all managers within MBPG and many of the workers that would be impacted by the changes. Also, the intern sponsor was made aware of the significance of making sure that the future of the project be understood and get the support from the management team that there will be a person dedicated to bringing this system back to where the DNA enters the facility.

7.3 Maintaining a Learning Organization

The long-term vision of the management of the operations organization at the Broad Institute is to leverage the education level, passion and innovativeness that their workers have to improve genome sequencing. The difference in the goals of this versus those of most manufacturing organizations is that the Broad wants to make these changes to provide people with the tools and abilities to make breakthrough changes in their processes, not incremental improvements.

This goal is unique about the Institute and provides many challenges when attempting to implement lean manufacturing systems. These challenges happen because the goals of the systems are not fully aligned, and alignment is very important for the success of organizational change. This misalignment is recognized by the management, and that is why many of the initiatives that they have been taking are through interns that work in the area, versus through a consulting agency that analyzes the Institute and recommends changes that are necessary.

As on manager at the Broad Institute stated, “I do not want Six Sigma and lean techniques to infiltrate the system such that their jobs become completely focused on a few quality and productivity metrics. I want these things implemented in a way that improves the flexibility of the operations and the workers. I want the workers to use the tools to innovate their jobs.”

Due to this strategic vision, management believes that they must be very careful about how they implement these changes and who they use to do it. They do not want someone to create an implementation that focuses on the Institute as if it were a major manufacturer such as a large automotive company. They want people who are familiar with academic environments, who will take the time to learn the production processes and understand the basics of the science behind DNA sequencing, and those who will tailor the needs of the organization to the processes being implemented.

This method of improvement, through several interns that are trained not only in manufacturing techniques, but also have a strong enough background in science to quickly come up-to-speed with the complexities of the process is important. It is also important for the Institute to bring in people that look past unique pieces of the Institute
or inefficiencies that are in place from the not for profit status of the Institute. The unique structure and goals make it a different place for implementation and the people behind it must be carefully screened as not to harm the goal.

The Institute has focused on hiring interns for long periods of time, usually one year at a time, and brought them into different groups. Groups that were thought to need more exposure to the techniques, like MBPG, have had several interns in succeeding years. This has been a strong attempt to bring about a culture shift in the organization and exposure to the workers through a series of successful projects.

That is one area of continuous learning that is positive for the Institute. Despite the vision, strategy, and resources that are used to meet this vision, others have not been focused on with the same level of detail.

For example, there is a high rate of turnover within the MBPG and there is also a strong difference in the skill level required for each of the processes within the group. Given that, the group does not often cross train their employees or encourage employees to apply for open positions within the organization or across different groups. Although there have been several people that have advanced within MBPG and other groups, the number of people versus the number of opportunities is very low. Additionally, there are few people within the group that can do multiples jobs.

These organizational choices have numerous consequences. First, it is believed that at least one of the major misregistrations could have been due to managers and inexperienced ligators helping the transformers with production. The reason that the manager was working on transformation was that none of the other workers in the area had any experience with transformation, but production required many people to help out.

Second, the group loses flexibility to meet variation in staffing and production. For example, the pilot for the sample tracking process was pushed off for one extra week because they did not have another person in the group who would be able to cover a vacation of a transformer and were worried about meeting production goals. The inability to respond to changes in production requirements, such as making new ligations and doing more transformations when the cell issue occurred can all be avoided if people learn multiple job functions. To make sure that people keep the work fresh in their minds, there should be a requirement of a certain number of hours each month that must be spent doing another process.

Finally, there is little room for people to develop to a higher level of understanding for the process. This prevents those at the lower levels of educational requirements from learning how the process works at a more in depth level. Although full understanding is not necessary, better understanding of the process and how specific roles fit into the process gives workers a framework for innovation. This will allow them to bring multiple experiences to their work and draw on the lessons that they have learned from each. Having workers with a myriad of experiences has proven to be valuable enough that many companies have created roles that force a portion of incoming workers to go
through rotation programs. Thirty of the Fortune 50 companies advertise rotation programs for employee development. These companies include; Intel, United Technologies, Johnson and Johnson, General Electric, Ford Motor Company and many others. It is understood that the Broad Institute has a unique culture and vision, in addition to being a not for profit organization, but the sheer number of companies that are considered to be successful that have these programs supports the claim that people find it valuable to have workers with a broad background.

Although the Institute is focusing on providing continuous learning for their employees by bringing in long-term interns to help change the culture and expose people to opportunities in collaboration with implementing systems like Six Sigma. There are other things that the Institute can do to improve the innovativeness and knowledge base of their employees. These include;

- Cross training
- Encouraging internal job growth, initially outside of their background
- Allowing for rotations in different groups and with different assignments

These will encourage employee development and align with the strategic vision by leveraging the skill of the employees to improve the quality and lower the cost of genome sequencing.

7.4 Conclusion

This chapter highlighted the difficulty an organization has to sustain an initiative, especially those driven by a temporary workforce. To sustain improvements that have been made and strive for future innovation and efficiency, initiatives must be driven by workers within the organization. These workers must believe in the initiative and understand how their work on the initiative can help the organization and their personal career growth.

Innovative behavior can be encouraged by changing Institute policies on cross training and employee development. Encouraging workers to expand their breadth of knowledge, depth of knowledge, and their ability to take calculated risks to improve processes can pay off significantly to the organization.
8. Conclusions

The Broad Institute is a growing organization that has a mission to provide genomic sequencing data to combat the fundamental issues that cause disease. The Institute is working toward the vision by developing capabilities in their research groups and by cooperating with other research organizations. They have also developed significant capabilities in the area of genomic sequencing operations. The capabilities have been grown through a significant investment of time and money into the Institute.

Although the goals of the organization are to combat disease, the majority of the way the goals are communicated is through the breakthrough research that happens at the Institute. This research is important to the success of the Broad as a whole, but it should be noted that the research would not be capable without the execution of hundreds of people in the genomic sequencing operations group performing the tasks necessary to obtain sequencing data.

Due to this, there are some recommendations that can be made to help the operations organization continue to improve and develop manufacturing capabilities in addition to communicating results of the Institute in terms of what the operations group relates to.

8.1 Developing Lean Manufacturing Capabilities

To continue to develop lean manufacturing capability within the sequencing platform, the organization can do a number of things;

- Hire a full-time employee who is trained in the techniques and have them lead projects and manage future interns.
- Institute Kaizen events within different areas of the organization to get groups of people exposed to the power that these projects can have if conducted correctly.
- Align the performance goals of the workers with improvement of their jobs, which can be done during a balanced scorecard implementation.

It is difficult to make large changes within the lean capabilities of an organization, especially without personnel dedicated to making the improvements and changing the culture within the organization. Internships help expose many different areas to the toolsets associated with improving these capabilities, but there is little consistency between internships on a year over year basis. This makes it difficult to make significant cultural changes. It will be better to have a dedicated person that develops credibility within the organization and constantly interacts with multiple areas within the sequencing platform.

Many corporations, such as Intel, Boeing, and Genentech use Kaizen events, or short project teams that attack a problem over a one or two week period of time. These projects usually result in significant improvements to the setup of the area and allow for multiple people to be involved in these events. Not only do the events do a great job of improving processes and increasing efficient use of work time, but they also involve workers in the thought processes associated with understanding what can be done to
improve an area and how one can analyze proposed changes. This would be an effective way for one full-time person to begin working with people across different areas of the organization and have a large impact by leading teams who implement the changes.

Finally, the goals of the workers are not aligned with the goals of continuous improvement or that of the Institute. Many of the workers feel that they have incentives to “not mess up”, rather than contribute to process improvements based on their knowledge of the process. One way to understand how to do this would be to implement a balanced scorecard system across the operations organization that matches with the goals of the Broad Institute as a whole. As this system funnels down to each group, the groups can develop metrics that support the Institute goals, but also encourage them to improve their skills in ways that align with the organization.

8.2 Educating Employees about Organizational Goals

It is very important to align the workers with the goals of the organization. The use of a balanced scorecard will help align the workers to this goal and understand how their performance on daily tasks has an impact to the Institute. Currently, the workers have little visibility into what is going on at a high level of the organization and do not often get an update on it. This year, the manager of the MBPG gave a great presentation at the end of the year that showed the workers what work they did and how it aligned to the goals of the Broad Institute for the year. The presentation highlighted accomplishments, challenges, and growth of the group over the span of the previous 12 months. It also highlighted the upcoming challenges and grants that the Institute was hoping to obtain and the reasons behind wanting those grants. This was a fantastic way to align the organization, but the workers were less aware of their contributions during the span of the year.

Presentations like that are not needed on a weekly or monthly basis, but insight into the data that was shown during the presentation would be very helpful to be posted in the area and standard intervals, such as monthly. This would allow the workers to understand if they are ahead or behind the goals of the year. They would also understand the reasons why they are working on the specific organisms that they are. Overall, it would help them gain a better connection between their work and the organization.

8.3 Final Recommendations

Final recommendations to the Broad Institute and the Molecular Biology Production Group would be to continue to improve processes that they are working on. Since many of the workers have been exposed to numerous organization, problem solving, and lean manufacturing techniques, it would be helpful to take the workers to other organizations that perform similar work. It is much easier to see the inefficiencies within ones own organization after visiting other organizations that do things differently. This perspective allows for the workers to gain a better understanding about how other environments work.
Ideally, the workers would be able to visit another genome sequencing operation so that they would have a direct comparison of their jobs to those that others do. When doing this, it would be helpful to give them an idea how their metrics compare. If the workers cannot visit a genome sequencing operation, there are numerous biotech and pharmaceutical companies in the Boston metropolitan area that would be happy to show the workers how things are done in a corporate manufacturing environment. This new perspective will allow the workers to begin seeing the organization that surrounds them in a different light.

The Institute should continue to bring people into the organization with an outside perspective. It is very valuable to have people that represent a different mindset and the Broad has done a fantastic job of doing this and rotating them through the organization. Once these interns are brought in, it would be helpful to have a common forum that the workers attend. This will give both the intern and the full-time workers at the Institute consistency between the internships, even if they are in different groups than the previous interns were.

Finally, invest time that the workers save through improvement projects into allowing them to make other improvements to the process or train in different areas. It is unfortunate that there is a deep stratification of worker roles within the MBPG. This feeling will make workers not want to learn the roles of others that are considered to have lower skill positions, even though they may need help for production reasons sometimes. The investment in training, like Six Sigma courses, is a fantastic long term investment in the employees and can pay significant returns if used wisely. In order to do this, the workers should begin finding productive projects to use their time wisely in between periods of incubation or biological reaction processes.

The Broad Institute is an organization that is gifted with highly skilled employees, a wonderful vision and worker dedication. It has invested significant resources to improve the capabilities of the genomic sequencing operations and should continue to do so. Better communication with the employees and providing a framework for the workers to take risks, to challenge their hypotheses, and improve processes within the Institute will be valuable to the future productivity improvements of the organization. This can be done by benchmarking other organizations and creating a framework for making changes and rewarding successful projects.
9. Author’s Biography

The author was raised in Des Moines, Iowa and graduated from Northwestern University with a bachelor’s degree in Materials Science and Engineering in 2001. During his undergraduate career, he worked at Mammoth Corporation and Masonite Corporation in research and development.

After graduation, the author worked at Intel Corporation for three years. During his time at Intel, he was responsible for large portions of the manufacturing process during third shift for the launch of the Pentium 4 and Centrino processors. He also developed capacity models and managed process improvement projects across the 0.13 micron process technology fabrication facilities. He left Intel to attend the Leaders for Manufacturing (LFM) program at MIT Sloan, where he received an MBA and a SM in Materials Science and Engineering.

In August, 2006, the author will begin working for Genentech in South San Francisco, California. The author can be reached at his permanent e-mail address - dapenake@sloan.mit.edu.
10. Reference

1 www.ornl.gov/sci/techresources/Human_Genome/home.shtml

2 http://www.broad.mit.edu/mammals/

3 www.broad.mit.edu

4 http://www.genome.gov/Pages/About/NACHGR/May2005NACHGRAgenda/ReportoftheworkingGrouponBiomedicalTechnology.pdf

5 http://www.alzheimers.org/unraveling/images/large/DNA-HIGH.jpg

6 http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/R/RecombinantDNA.html

7 http://genetics.biol.ttu.edu/genetics/Lecture/red27.html

8 Data obtained through Broad Institute cost information database (2005)

9 Nicol, Robert, Presentation at the Broad Institute, October 12, 2004

10 Sample misregistration occurs when a sample is processed in sequencing operations under an incorrect organism identification number. For example, the DNA may be believed to be a bushbaby by the sample identification number, but the DNA is actually that of a mouse.

11 Data compiled through analysis of transformation processes.

12 http://www.genome.gov/15015208

13 Adapted from Vokoun, Matthew R., “Operations Capability Improvement of a Molecular Biology Laboratory in a High Throughput Genome Sequencing Center”

14 Some information has been removed, but the table is used for illustration purposes.

15 Data compiled based on information available through materials cost information at the Broad Institute.
Data was collected and analyzed after pilot. All errors were fixed prior to sending material to downstream processes.

Figure illustrates the addition of barcode tubes, elimination of plate steps and prevention of sample re-entry points.

Cell colony counts below 500 or above 4500 are generally accounted as failures.

San, Ka-Yiu, Bennett, George N., “Expression Systems for DNA Processes”, Enzymes, Molecular Biology and Biotechnology: a Comprehensive desk reference

http://en.mimi.hu/biology/ethanol.html


Christopher, Raso, Zeigler, Baldwin, Enzymes, Molecular Biology and Biotechnology: A comprehensive desk reference


Eisenthal, Robert, “Enzyme Assays”, Molecular Biology and Biotechnology: A comprehensive desk reference

Eisenthal, Robert, “Enzyme Assays”, Molecular Biology and Biotechnology: A comprehensive desk reference


33 San, Ka-Yiu, Bennett, George N., “Expression Systems for DNA Processes”, Enzymes, Molecular Biology and Biotechnology: a Comprehensive desk reference


37 Baum, C. Transfection - Encyclopedia of Molecular Biology, Volume 4, Wiley Biotechnology Encyclopedias, pp. 2596 – 2600


44 Vokoun, Matthew R., “Operations Capability Improvement of a Molecular Biology Laboratory in a High Throughput Genome Sequencing Center”

45 Ligation production data over a period of 1 year within the MBPG. The data depicted is the mean and standard deviation of cell colony counts (white counts) from 4 libraries created with each ligation.

46 Data compiled on 6/16/2005 at the Broad Institute

Data collected at the Broad Institute on various dates in November 2005.

This is a rule of thumb at the institute and is not believed to have data that supports the conclusion.

Analysis was done in collaboration with Scott Couzens, LFM ‘06

E-labnotebook is an online system that is intended to allow researchers at the Institute to share data with one another with intranet software. This system is also intended to provide a record of research an processes to prevent total operations loss due to a catastrophe.


Klein, Jan, “True Change”, John Wiley and Sons, 2004

Vokoun, Mathew 2004

Nicol, Robert, Director of Sequencing Operations, The Broad Institute

http://money.cnn.com/magazines/fortune/fortune500/


