Clear Labs

An Introduction to Next-Generation Sequencing for Food Safety Labs

Table of Contents

Chapter 1
An Abridged History of Food Safety Technology

In 1885, Theodor Escherich discovered *E. coli* in the human colon. In that same year, Theobald Smith and David E. Salmon found the first strain of *Salmonella*. Since then, our knowledge of foodborne pathogens has increased drastically, but detection methods have been slow to change. Cell culturing, a methodology with its roots in the nineteenth century, still remains the gold standard for confirmation in the food safety industry.

Of course, culturing methods have evolved greatly since 1885, when Wilhelm Roux created his first method of cell culture and demonstrated that it's possible to maintain living cells in a saline buffer. Nevertheless, cell cultures still present challenges for food safety professionals. Foremost among these challenges is the method's speed. Growing cultures requires several days to a week to generate confirmed and accurate test results. Food manufacturers and labs require faster turnaround times so that they can release inventory faster.

Philip Perlmann and Eva Engvall found a solution to the question of speed in 1971. Their work led to the commercial development of an enzyme-linked immunoabsorbant assay (ELISA) technology. ELISA is a rapid substance detection method that can identify a specific protein, like an allergen, in a cell by binding an antibody to a specific antigen and creating a colorimetric or fluorescent reaction. It is less effective in food testing for cooked products, in which the protein molecules may be broken down and the allergens thus are no longer detectable.

About a decade later, in the 1980s, Kary B. Mullis invented the polymerase chain reaction (PCR). PCR was a major breakthrough because it allowed scientists to make numerous copies of a small amount of DNA. As a result, detection sensitivity and specificity improved, and scientists were able to study specific target regions of DNA.

Needless to say, it was an exciting time when commercial real-time PCR instruments reached the market in the mid-1990s. To date, PCR testing remains the industry standard for pathogen detection despite the technique's shortcomings, including high rates of false negatives and positives, as well as binary yes/no answers.

By using PCR, food safety professionals were able to answer the question: "Does my sample contain a specific microbe?" Starting in the 1990s, the scientific community began to ask a different question: "What is the microbial makeup of my sample?" PCR cannot generate the data needed to answer this question. The technology needed is next-generation sequencing (NGS), a rapid and high-throughput method of sequencing DNA.

NGS was developed from the requirements of the Human Genome Project, started in 1990 by the U.S. government, in conjunction with several other countries. The goal was to map all of the genes of the human genome. The process wasn't completed until 2003, and the cost was $2.7 billion. Since then, the price and turnaround time for mapping genomes have dropped significantly.

While the human genome project was in the process of being completed, scientists began to map the genomes of foodborne pathogens. In 1997, in the wake of the Jack in the Box *E. Coli* outbreak, the *E. coli* genome was sequenced by Frederick R. Blattner and his fellow researchers. In turn, the food safety industry was introduced to DNA sequencing.

Chapter 2
NGS: A Breakthrough Approach to Food Safety

Mapping an entire genome is just one application of NGS. As we will discuss in upcoming chapters, there are many applications for next-generation sequencing. But for now, know that NGS is an umbrella descriptor for sequencing millions of DNA data points at the same time.

With so much data at their fingertips, scientists can learn more about food and environmental microbial communities than ever before. For example, by using NGS platforms, companies can conduct accurate persistent/transient studies, determine product freshness and quality through microbiome studies, verify suppliers through authenticity tests, launch traceability studies, and decrease the rate of false negatives and positives of their pathogen detection studies.

But that's not all. NGS platforms consistently outperform legacy technologies like PCR. In real-world applications, NGS platforms have proven their ability to deliver highly accurate results (i.e. >99.9% accuracy), dramatically lowering false negatives and positives.

What is more, NGS platforms are more flexible than legacy solutions like PCR. They allow food safety professionals to drill down to the desired level of detail. For example, when using an NGS platform for pathogen detection, food safety professionals can decide if they want to simply run a detection test for Salmonella at the genus level, or they can decide if they want to also run serotyping tests and which serotypes they want to test for.

	PCR	NGS
Accuracy	Higher rates of false positives and negatives	>99.9% accuracy, resulting in dramatically lower rates of false positives and negatives
Pathogen Profiling	Expensive and slow speciation with limited serotyping capabilities	Speciation, serotyping, and strain identification are faster, more inexpensive, and carried out simultaneously
Throughput	Screen for one pathogen at a time	Screen for an array of pathogens in one test
Insights	Binary yes/no answers	Professionals can choose the depth of information that they want

From a purely financial perspective, NGS platforms are equally impressive. While the human genome project cost $2.7 billion, NGS platforms are now within reach of any organization. In fact, for routine speciation testing, the price is the same as a PCR test, and as a result, NGS is poised to replace PCR in routine food safety testing.

As an added financial bonus, NGS platforms reduce costs by simplifying the workflow for pathogen screening and serotyping. Right now, most food manufacturers and labs use PCR or immunoassays to test for presumptive positives, which takes about 24 to 36 hours. Afterwards, confirmation tests must be performed, which can take another 24 hours to 5 days, depending on the methods used.

With NGS, that amount of time is shortened. Food manufacturers and labs can perform detection and serotype tests roughly 24 hours after the primary enrichment, and since NGS tests are >99.9% accurate, food safety professionals can release their products with confidence, reducing costly holding times.

We'll further discuss the ROI of NGS in the sixth chapter of this book, but for the time being, it suffices to say that NGS platforms deliver a substantial price/performance advantage, making it one of the breakthrough technologies of the food safety industry.

Chapter 3
Three Key Applications of NGS and How to Use Them

It can be hard to wrap our minds around NGS. It's a test that produces millions of data points. So what? How are millions of data points useful for food scientists?

To answer that question, we must remember that NGS is an umbrella term. Think of NGS like a smartphone. iPhone users, for instance, rely on different applications for messaging, for their finances, for entertainment, and so on. The same is true of NGS. There are several applications of NGS, and each application serves a different purpose.

Let's look at three of the most applicable applications for food safety: WGS, shotgun sequencing, and targeted NGS. As we look at those, we will discuss what they are and when they should be used in food safety programs.

	1 WGS	2 Shotgun Metagenomics	3 Targeted NGS
Type of DNA Sequencing	Non-targeted	Non-targeted	Targeted
Testing Source	Pure isolate-based	Isolate independent	Isolate independent
Turnaround Time	Several days	Less than a day to several days	Less than a day
Cost	Expensive	Expensive	Same cost as PCR
Sample Source	Limited to a pure isolate	Applicable to complex cultures	Applicable to complex cultures
Detection Sensitivity	Sensitivity not applicable	Variable sensitivity	Highly sensitive
Example Use Case	Traceback analyses	Broad investigative applications	Routine pathogen testing

1 Whole-Genome Sequencing: Great for Traceability Studies

To start, let's take a look at WGS. Most food safety professionals have probably heard of WGS, as it continues to grow in prominence. In short, WGS is the process of determining the complete DNA sequence of an organism. For a long time, only researchers and academics used WGS because it was time-intensive, costly, and needs special expertise to analyze the generated sequences. However, upon the introduction of NGS platforms, the cost of WGS has come down significantly. As a result, WGS has become routine in the clinical and food safety spaces.

WGS is a powerful tool for situations in which a lab needs detailed genetic characterization of a pure sample. By doing WGS of a given sample, the most detailed genetic footprint of that sample is created. This could include up to several million genetic building blocks per sample, for a bacterial sample. These genetic building blocks are then stitched together to create larger genetic scaffolds and then compared to a reference database to identify the microorganism being studied. Computer algorithms can notice the slightest differences between the DNA of the samples and that of the microorganisms in the database.

In the food safety space, WGS is not a good fit for routine pathogen testing because it is costly and slow and requires a pure culture. Instead, traceback analyses are prime candidates for WGS because this approach to NGS gives the resolution required to differentiate between strains and pinpoint where a food sample picked up, say, Salmonella Kentucky strain sequence type (ST) 198. Moreover, WGS is ideal for traceability studies because such studies rely on pure samples, which are prerequisites for WGS studies.

Traceback analyses aren't the only use cases for WGS. Food safety scientists can use WGS whenever they need detailed characterization of a given genome and its functional attributes. For example, they can leverage WGS to determine which antibiotics an animal is resistant to and which nutrients they can metabolize.

2 Shotgun Metagenomics: Ideal for Broad Investigative Applications

While WGS has decreased in cost, the application's price point and turnaround time are not ideal for routine testing. What is more, the test requires pure cultures, which can be limiting. On the other hand, shotgun metagenomics, though expensive, does not have the same limitations on cultures. It can be run on pure or mixed microbial populations.

In shotgun metagenomics, DNA is randomly shredded into many small fragments that can be sequenced individually. Then, the sequences of these fragments are reassembled into their order, ultimately producing the complete

sequence. Shotgun sequencing is unbiased and can generate unparalleled details about the identity of all genetic material within a biological sample. WGS is usually achieved by applying a shotgun approach to a pure sample. If the sample is complex, then applying shotgun sequencing generates shotgun metagenomics data.

A shotgun metagenomics approach to NGS is great for broad investigative projects like environmental monitoring studies, where food safety professionals need a complete view of the microorganisms in their lab and their functional attributes. Take, for example, biological monitoring programs. A shotgun approach can be used to analyze bulk environmental water or soil samples looking for the presence or absence and function of indicator organisms, and machine-learning software can use the data to reconstruct the networks of ecological interactions. That way, food safety professionals can identify and minimize the risk of contamination.

There are some complicating factors with a shotgun approach. Imagine an environment with communities of bacteria with 10,000 different strains. However, you only want detailed characterization of Listeria monocytogenes. It can be difficult to arrive at the right depth of information. Additionally, if the sample has a very high degree of complexity, the amount of data needed to confidently characterize the target microorganism could be very high, and by extension, the endeavor can become cost inefficient.

3 Targeted NGS: Built for Routine Pathogen Testing

So far, we've discussed two non-targeted NGS applications. Now, let's take a look at targeted NGS. Whereas WGS looks at the entire genome, targeted NGS looks at many loci/genetic markers within a broader pool of genetic materials. While this might sound like a limited PCR test, it's different in a few fundamental ways. PCR-based tests only look at one genetic marker (or, in case of multiplex PCR, at a couple of genetic markers) rather than multiple genetic markers as is done in targeted NGS. As a result, the accuracy of PCR tests is lower than that of targeted NGS.

The other fundamental difference between PCR and a targeted sequencing-based approach is that PCR can only flag the presence and absence of a genetic marker. On the other hand, targeted DNA sequencing can read every single building block of a targeted region to make sure that the test amplified the correct target. Then, it can use algorithms to compare the sample's building blocks with a reference database and confirm whether the target matches the DNA sequences of, say, Salmonella Saintpaul and not Salmonella Derby.

Another differentiation factor between PCR and targeted sequencing is the amount of information collected. Targeted sequencing can be utilized to increase the amount of information collected from a single bit (in case of PCR) by sev-

eral orders of magnitude. Such high information content, beyond guaranteeing a superior performance, when combined with AI and machine learning can generate new layers of insight, like predictive analytics and analyzing trends that were not possible before. For example, targeted sequencing has the potential to lend more accuracy to shelf-life predictions.

To be sure, targeted NGS and PCR do share some commonalities. For starters, the turnaround time for detection is the same with both PCR and targeted NGS. In addition, both PCR and targeted NGS are far cheaper than both WGS and shotgun metagenomics. In fact, PCR and targeted NGS tests can have similar costs.

But with targeted NGS, scientists can do much more than they can do with a PCR test. Targeted NGS platforms are ideal for routine pathogen testing. Because of their targeted nature, these platforms can be fine-tuned to be very sensitive, detecting a needle in a haystack. Or, to put it in scientific terms, targeted NGS platforms can detect a specific serotype in a mixed sample, a pure sample, an individual colony, a liquid sample, or a solid sample. As a result, targeted NGS platforms are ideal for any application where accuracy is of the utmost importance.

Due to the superior accuracy of targeted NGS solutions, food manufacturers and labs can expect fewer ambiguous results that require lengthy confirmation steps. In that sense, targeted NGS testing can save companies both time and money by reducing their inventory hold times inherent to the confirmation process.

Chapter 4
Deeper Dive: Why WGS Isn't the Right Tool for Routine Pathogen Testing

Most food safety professionals have heard of WGS, and as a result, a misconception has run amok in the food safety industry. You've likely heard the terms "whole genome sequencing" and "next generation sequencing" used interchangeably. However, while the technologies are related, they're not the same.

WGS has proved useful for food safety professionals, becoming instrumental for traceback analysis of foodborne illnesses and aiding in spoilage and shelf-life analyses. Compared to conventional methods, WGS has been able to better detect small outbreaks, and we can obtain strain-level information about antibiotic resistance, virulence factors, and gene annotation from WGS data.

That said, WGS is mainly an investigative tool, and it is not the right approach for routine pathogen testing for a

variety of reasons. For example, WGS is still far too expensive for any type of daily or weekly testing. Even the most efficient labs spend hundreds of dollars per sample to do WGS. These labs also need to be staffed by highly trained professionals, as WGS requires special expertise to prepare sequencing libraries. These procedures can only be automated on a small scale, making the overall process demanding and costly.

Granted, WGS no longer takes 13 years to conduct, as it did when the human genome was first mapped. Nevertheless, it still is too slow for most food manufacturers.This is because the technology requires a clonal isolate to collect the necessary data, which takes time and intensive microbiological work to generate, on top of the time required for sequencing. Test-and-hold programs need results quickly. Manufacturers cannot wait several days to a week for results.

With WGS approaches, food samples are also complex to the point of frustration. One food sample alone might contain several serovars of a bacteria. But WGS can only analyze one serovar at a time, thus giving you a limited view of your food sample.

Finally, WGS generates a lot of data points. Some food companies may not feel comfortable analyzing all that data, while others may not feel the need to have so much data at their fingertips.

As we discussed in Chapter 3, targeted NGS is better suited for routine pathogen testing. To recap, here are the main reasons why:

- **NGS is cheaper and faster than WGS while still maintaining >99.9% accuracy**
- **It can work on clonal isolates or on mixed populations, something that WGS cannot do**
- **The results of a targeted NGS test are much easier to read**

Chapter 5
NGS Is More Than a Diagnostics Test

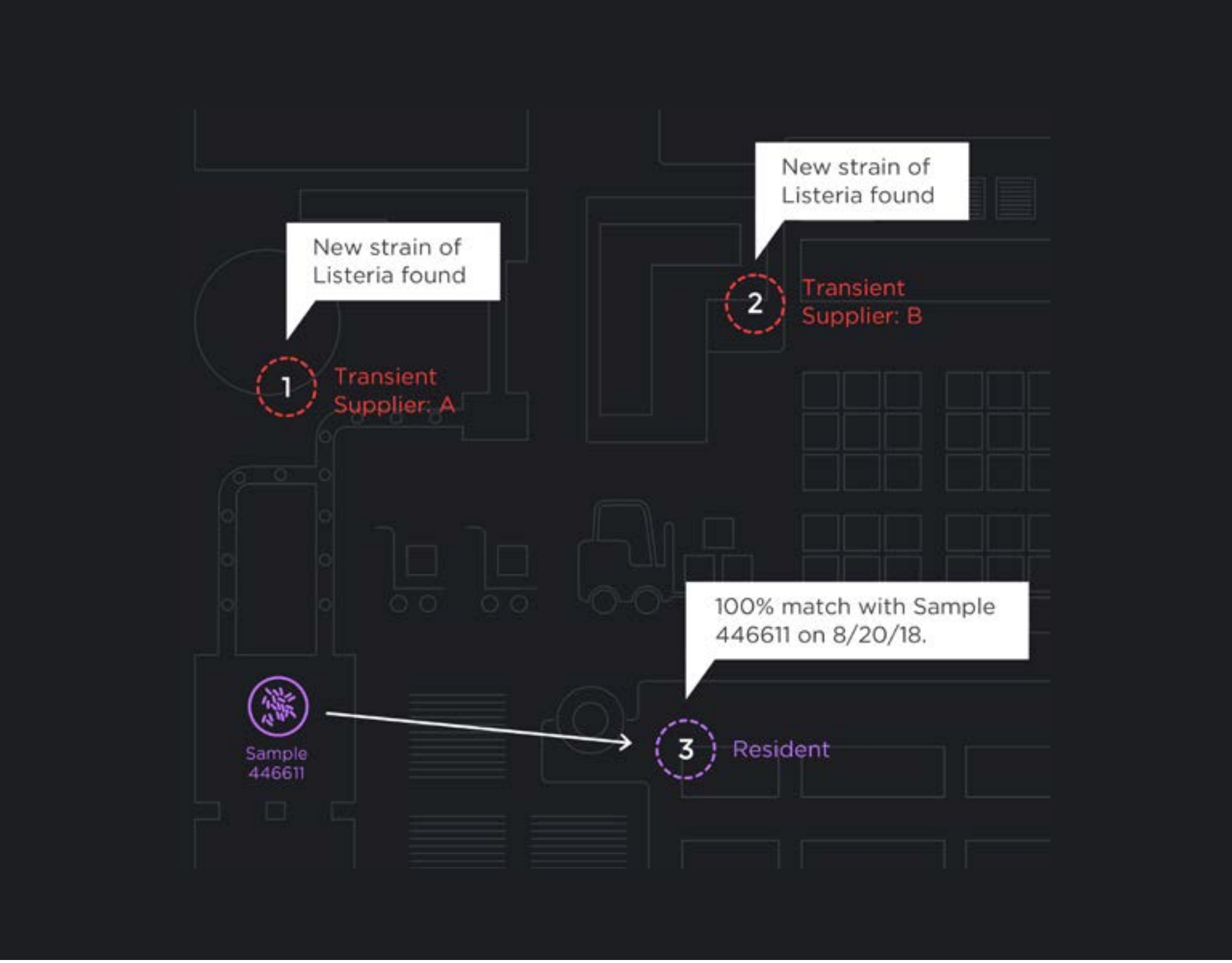
Mention NGS to those with a cursory knowledge of the subject, and it often conjures comparisons to PCR. To a degree, the comparison to PCR is apt because both NGS and PCR can be used for routine pathogen testing. But NGS can also do so much more. In this chapter, we'll explore some of the ways in which NGS platforms are more than a mere diagnostics test.

Platform Consolidation

The food safety technology landscape is fragmented. Most companies use one platform for routine pathogen detection, another for serotyping, another for persistent/transient studies, and yet another for environmental mapping. But with NGS, platform consolidation is possible, thus reducing costs and increasing efficiency.

Take the example of Listeria. An NGS platform like Clear Safety combines speciation testing, persistent/transient technology, and environmental mapping technology.

By combining three platforms in one, NGS platforms provide food safety professionals new insights. For example, Clear Safety links persistent/transient data to facility maps, enabling users to visualize the migration of a pathogen throughout a facility and to link transient pathogens to suppliers when desired.



Robotics

Once only within reach of the most well-funded labs with specialized technicians on staff, automated library preparation is now accessible to nearly any food safety lab. Automation has become more affordable and much simpler to operate.

To streamline their processes and improve turnaround times, many labs are turning to robotic liquid handlers. This is a robot that dispenses a selected quantity of reagent, samples, or other liquid to a designated container. (You can watch an example here.) By relying on a robot, food safety labs can increase reproducibility with reduced hands-on time and technician error.

Robots might sound intimidating, but top-of-the-line systems will provide step-by-step instructions, thus reducing potential errors. Plus, the system will display a summary screen that acts as a final check, ensuring that anyone in the lab can use the instrument.

Automating Bioinformatics Workflows

Beyond robotics, NGS brings automation to our existing bioinformatic workflows through artificial intelligence and machine learning. NGS provides an enormous amount of data, much of which goes unused in food safety applications today. The next revolution in food safety is in automating data-science operations that can leverage this data.

Automating bioinformatic workflows will dramatically increase the industry's ability to analyze enormous bodies of data and identify macro-level trends. Imagine the insights you could gain when you combine trillions of genomic data points from each phase in the food safety testing process — from routine pathogen testing to environmental monitoring to serotyping. You can create better environmental maps. You can build models to predict shelf life. You can run predictive risk assessments.

Don't worry. Your lab won't need to hire hundreds of bioinformaticians to crunch the numbers. A good NGS platform will build the algorithms that run in the background and generate reports that anyone in your lab can read and understand.

Chapter 6
The Economic Argument for NGS in the Food Safety Lab

Working in a cost center can be challenging at best and soul-sucking at worst. Many food safety professionals have heard their departments referred to as "drains on profits" or simply as "overhead," and they've felt pressure to reduce costs.

But how? There are only so many levers one can pull. Technology vendors have limits on their price negotiations. Companies still need to pay their lab workers. Plus, there are the costs associated with inventory holding times and diverted or destroyed products.

That's where next-generation sequencing (NGS) can help. By abandoning legacy technologies like PCR and embracing NGS, food safety labs can transform a cost center into a margin contributor and make positive contributions to the bottom line.

Reduce Costs Associated with False Positives

False negatives grab headlines in the press, but false positives are thorns in the sides of any food manufacturer. When a false positive presents itself, companies have to absorb the costs associated with destroyed, diverted, or re-worked products. On top of that, federal agencies require equipment examinations, intensified sampling and testing, reviews of procedures and records, and so on.

All this adds up. In 2007, an incorrect association of tomatoes with Salmonella Saintpaul resulted in a nationwide recall that cost growers and packers more than $30 million.

Needless to say, life would be easier for food manufacturers if they could reduce the number of false positives they encounter. That's a key benefit of NGS. With greater than 99.9% accuracy, NGS platforms dramatically reduce the rates of false positives and all the costs associated with them.

Reduce Inventory Holding Time

Gaining control of inventory is a pain point for companies in the food and beverage industry Margins are slim, products are often perishable, and cross-contamination can happen if foods are stored in proximity.

That's why many executives are eager to ship product quickly.

In the infamous case of the Peanut Corporation of America (PCA) from 2008 and 2009, Stewart Parnell allegedly instructed his employees to "just ship it" when they wanted to know what to do when test results came back positive for *Salmonella*. The result? A four-year investigation by the U.S. government and $144 million in economic losses.

Of course, the PCA example is an extreme case. When following standard protocols, companies typically hold inventory while they await confirmation and serotyping results after positive tests.

What are the daily inventory holding costs for your company? Take that number and multiply it by the turnaround time for your confirmatory tests. If you consider that confirmation can take up to five days, that adds up quickly. Granted, waiting for confirmation is certainly cheaper than the $144 million that PCA faced or the millions in hard costs and brand destruction a company faces if a recall is required. Nevertheless, holding costs are unwanted expenses.

An NGS platform can help alleviate this pain point. Within roughly 24 hours, from sample to result, NGS-based testing can deliver not just speciation results, but serotyping results on positives as well. This condenses the amount of time needed to receive serotype information significantly.

What is more, since NGS platforms have greater than 99.9% accuracy, not only will the number of false positives be reduced; some companies may view NGS-based positive results deterministically. In other words, they will take a positive result from an NGS platform as a true positive. Instead of waiting for confirmation, they will immediately divert, rework, or destroy their products, thereby cutting their losses.

Reduce Hands-On Labor

Lab managers are always searching for ways to eliminate bottlenecks and save on labor costs. “Walk-away” time is of the utmost importance because it gives employees the opportunity to move beyond hands-on labor to accomplish other research and development activities.

That’s where NGS platforms enter the scene. A good NGS platform, purpose-built for food safety, will automate some of the menial tasks through robotic liquid handlers, which are geared for high-throughput testing. Because of these innovations in automation, employees do not need to intervene during key stages of the testing process and have more time to work on new projects.

That said, NGS technology’s contribution to the reduction of hands-on labor is not simply a question of automation. NGS platforms make an operational contribution, as well, because they are the ultimate consolidators of technology. Right now, speciation and serotyping are performed separately, requiring additional labor for both. With NGS platforms, however, companies can perform speciation and serotyping at the same time, thus eliminating costly, labor-intensive steps.

Similarly, NGS platforms can test for multiple pathogens at the same time, meaning lab technicians perform sample preparation only once to test for an array of pathogens at the same time, say, Listeria and Salmonella. In the past, employees might have to perform two separate sample preparations to test for both pathogens--only to get less accurate results.

Just think of the strategic initiatives that you can undertake if you free up employees!

Chapter 7
Introducing Clear Safety

It's time to nudge your executive team to think differently about food safety. It does not need to be seen as a "cost center." Instead, your food safety team can be an asset to the company. Sure, top-line growth is the primary concern of most companies, but reducing costs is a critical consideration. By switching to an NGS platform, your food safety program can be good for your company's bottom line.

Currently, there's only one platform that can help. That's Clear Safety. PCR, culturing, and antigen-based tests tell you whether a pathogen is present or absent. Clear Safety is much more. It is the first automated, intelligent NGS platform that's purpose-built for food safety testing. Clear Safety generates hundreds of millions of data points per analysis that can be used to learn significantly more than what today's pathogen screening platforms allow.

Features & Benefits

Pathogen Profiling Gives You the Info You Need, When You Need It

Clear Safety enables deep molecular characterization of pathogens, with just the amount of information needed for your safety program. You can choose to drill down to serotypes and strains based on the application need.

Multi-Target Analysis Makes Food Safety Cost-Effective

With Clear Safety, you can efficiently screen an array of pathogens, increasing throughput and reducing costs. Clear Safety enables you to run pathogen samples in the same or different matrices in parallel. Plus, you can save money by eliminating multiple pathogen tests per matrix or by volume batching.

Automation Curtails Errors and Inefficiencies

Clear Safety's automation guarantees reproducibility and reduces hands-on labor and error potential. Our automated workflow reduces hands-on time, technical error, and variability, and it provides more flexibility around shift schedules. Our barcoding system makes it easy to track samples and reduces data entry errors.

Superior Accuracy Sharply Reduces False Positives and Negatives

Clear Safety's accuracy sharply reduces false negatives and positives, curtailing recall risks, operational costs, holding time, and short-shipping penalties. With >99.9% accuracy, we give you the best screening assurance in the industry. You can stop chasing ghosts and start shipping product.

Fast Turnaround Time Allows You to React Quicker

Within 24 hours, you can find out serotype information and take action. You can eliminate inventory hold time waiting for a positive confirmation. With Clear Safety, you can react on a presumptive with >99.9% accuracy.

Enterprise Software & LIMS Simplifies Decisions

With our software and barcoding system, you can easily track samples from enrichment to answer. Our cloud-based solution gives you access to faster reporting and data analytics on samples and sites. What is more, our reports are easy to read, thus simplifying decisions for your team and enabling you to release product with confidence.

Want to Learn More?

Send us an email at inquiries@clearlabs.com, and we'd love to discuss how we can help you.

Appendix - Key Terms

Genome: The complete set of genes or genetic material in an organism.

WGS: Whole genome sequencing. WGS uses NGS platforms to look at the entire DNA of an organism. It is non-targeted, which means it is not necessary to know in advance what is being detected. In WGS, the entire genome is cut it into small regions, with adaptors attached to the fragments to sequence each piece in both directions. The generated sequences are then assembled into single long pieces of the whole genome. WGS produces sequences 30 times the size of the genome, providing redundancy that allows for a deeper analysis.

Shotgun Sequencing: DNA is randomly shredded into many small fragments that can be sequenced individually. Then, the sequences of these fragments are reassembled into their order, ultimately producing the complete sequence.

Shotgun Metagenomics: This is a specific type of shotgun sequencing. If the sample is complex (i.e. has multiple populations) as most environmental samples do, then applying shotgun sequencing generates shotgun metagenomics data.

PCR: Polymerase chain reaction. First described in 1985, PCR is a technique to amplify a segment of DNA and generate copies of a DNA sequence. The DNA sequences generated from PCR must be compared to specific, known pathogens. While it can identify pathogens at the species level, PCR cannot provide the strain of a pathogen due to the limited amount of sequencing information generated.

ELISA: Enzyme-linked immunosorbent assay. Developed in 1971, ELISA is a rapid substance detection method that can detect a specific protein, like an allergen, in a cell by binding antibody to a specific antigen and creating a color change. It is less effective in food testing for cooked products, in which the protein molecules may be broken down and the allergens thus no longer detectable.