CUSTOMER SPOTLIGHT

Harnessing the Power of Technology to Understand Biological Phenomena Begins in the Lab

A University of Michigan lab combines machine learning and high-throughput technologies to more rapidly develop scientific tools and advance knowledge.

Going to the dentist is a stressful experience for many people, especially when filling a cavity is required. An improved understanding of how to prevent dental caries (cavities) could offer some relief. However, research on the complex community of bacteria living on our teeth, specifically the cavity-causing bacterium *Streptococcus mutans*, can be labor-intensive and expensive.

The Jensen Lab at the University of Michigan uses laboratory automation, machine learning and functional genomics to tackle biological challenges with combinatorial structures. From left to right: Vasu Rao (MS student), Noelle Toong (Ph.D. student), Kurt Kostan (Ph.D. student), Paul Jensen (Principal Investigator), Ryan Wyllie (Assistant Research Scientist), Deepthi Suresh (Ph.D. student), and Noah Schmid (Ph.D. student). Not pictured are Benjamin David (Ph.D. student) and Brendan Brasch (MS student).

Fortunately, researchers at the University of Michigan are combining new, cost-effective technologies and machine learning to accelerate research on oral microbes. Thanks to their generalizable methods for building genetic engineering toolkits and efficiently measuring gene expression, it's now possible to explore the community of bacteria responsible for many dental diseases without a huge budget.

To learn more about this approach, we interviewed Dr. Ryan Wyllie, an assistant research scientist in the lab of Dr. Paul Jensen at the University of Michigan. The Jensen Lab focuses on "biological problems that are too large for either experiments or computation alone." Using automation and machine learning, Wyllie led a team to streamline the characterization of genetic processes within a group of bacteria called the oral streptococci, specifically the cavity-causing bacterium *S. mutans*.

Wyllie's goal was to develop a robust, high-throughput technology that could quickly and efficiently extract information about bacterial gene expression. The result was a new, low-cost transcriptomics technology called Splintlock-seq. Splintlock-seq converts the abundance of mRNA transcripts into DNA libraries compatible with next-generation sequencing (NGS). Splintlock-seq uses computationally designed, genome-wide padlock probes and the activity of SplintR ligase to increase the assay's specificity while reducing its cost.

A major challenge in the development of Splintlock-seq was ensuring that the relevant sub-protocols could be scaled and easily automated. To address this, the lab routinely used Beckman Coulter's RNAclean XP beads for simultaneous PCR cleanup and fragment size selection. The Emnetik System further increased the lab's genomic sample processing speeds while reducing the amount of hands-on time required. Both Beckman Coulter products streamlined the development of Splintlock-seq.

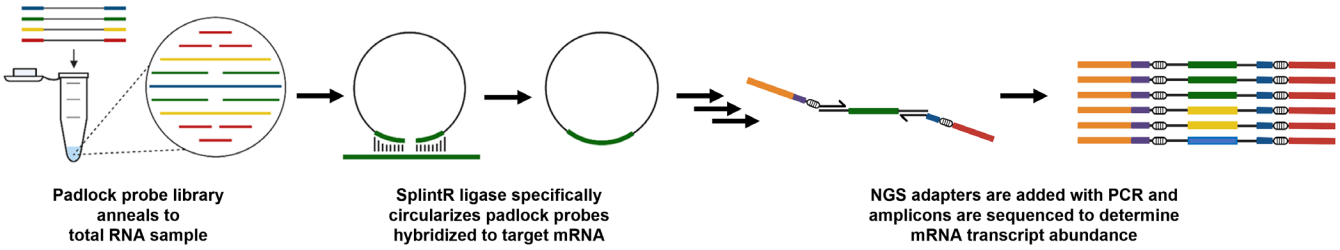

Figure 1. Overview of a Splintlock-seq library prep. In a highly specific manner, SplintR ligase converts ssDNA padlock probes hybridized on their mRNA target into circular ssDNA molecules. These circularized molecules are then subsequently amplified using primers which add adapters and indices for next-generation sequencing.

The multidimensional optimization of Splintlock-seq relied heavily upon statistical design of experiments (DOE), which in turn required many experiments. Size-selection with gel electrophoresis and column-based purification for final Splintlock-seq library cleanups would have been extremely cumbersome, limiting the scope and efficiency of relevant experiments. But the consistency and scalable nature of RNAclean XP bead-based cleanups enabled smooth execution.

During routine PCR cleanups with RNAclean XP beads, the overall yields were similar to silica-column based preps, but the consistently high purity of DNA isolated with RNAClean XP beads made them a significantly better option. PCR cleanups with the Emnetik reliably resulted in >80% yields with low batch to batch-to-batch variation, regardless of the user's familiarity with bead-based purification protocols.

While the work of streamlining the characterization of bacterial genetic processes is ongoing, the quality and quantity of data collected proves the value of combining machine learning with high-throughput technologies. In another example from the lab, Ph.D. student Benjamin David combined experiments and modeling to improve a key step in preparing genomic libraries, published in SLAS Technology[1].

Wyllie continues to be excited about using technology in the lab to further scientific exploration.

"I've always been fascinated by what bacteria are capable of and inspired by the potential of leveraging those capabilities to help humanity," he said. "We've observed bacteria that can survive on the *outside* of the International Space Station, bacteria that can "eat" steel, and bacteria that make all sorts of potentially therapeutic compounds. However, in order to truly harness those capabilities, we need tools to quickly and efficiently learn about gene function and gene expression. And that's where we come in."

1. Benjamin DM, Jensen PA. Improving an rRNA depletion protocol with statistical design of experiments. SLAS Technology. 2022. https://doi.org/10.1016/j.slast.2022.09.004

2023-GBL-EN-100818-v1