A Novel Engineered Transposase (TnX™) for Improved Performance in Genomic Applications

David Bays[1], Jonathan Vroom[2], Kalyan Krishnamoorthy[2], Aksiniya Petkova[2], Stella Huang, Ashley Silvia[1], Yanyan Liu[1], Rebecca Feeley[1], Zac Zwirko[1], Maura Costello[1], Jack T. Leonard[1], Gavin Rush[1], Mathew Miller[3], Joseph Mellor[1]

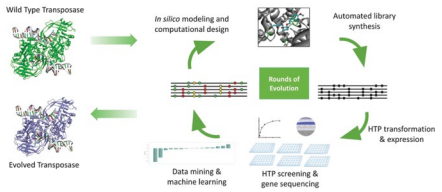
[1]seqWell Inc, R&D, Beverly, MA, USA. [2]Codexis, Protein Engineering, Redwood City, CA, USA. [3]Codexis, Life Science and Technologies and Applications, Redwood City, CA, USA.

Introduction

- TnX™ is a novel high-performance transposase, engineered for lower insertion-site bias, improved activity and robustness in library preparation using Codexis' CodeEvolver® platform.
- When compared to Tn5, human 15X WGS libraries made with TnX show significantly reduced insertion bias, leading to better uniformity of coverage and increased library complexity. We additionally show improved coverage evenness in our ExpressPlex™ rapid library kit.
- Combined with seqWell's streamlined multiplexed workflows, TnX enables high quality low bias library prep with less time and labor.

CodeEvolver® Technology

- Codexis' CodeEvolver® technology rapidly generates enzymes with desired properties through iterative rounds of enzyme optimization (at right).
- Starting with a novel artificial sequence with transposase activity, in silico modeling and computational methods were used to engineer large enzyme libraries with a variety of mutations.
- High throughput (HTP) expression and screening methods were used to develop a panel of top hits that were then further screened in seqWell's library prep workflows. This cycle was repeated until the desired performance specifications were achieved.

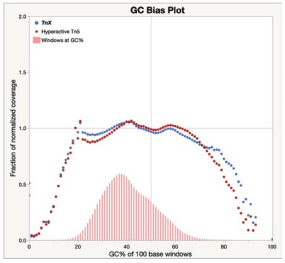


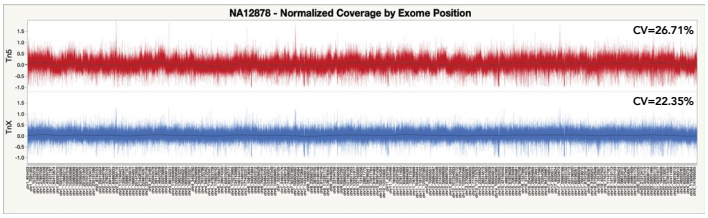
Performance in Human Whole Genome Sequencing

- Replicate WGS libraries were made using 100ng of human gDNA NA12878 (GIAB reference) in a high complexity library workflow using hyperactive Tn5 or novel TnX. The libraries were then sequenced to a depth of 620M reads per library on a NovaSeq X (~16X coverage).

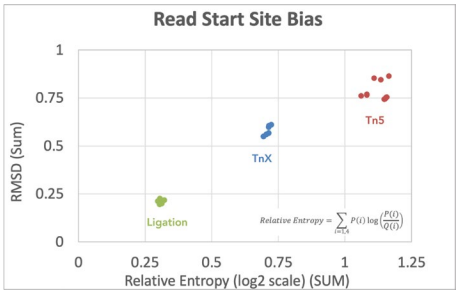
WGS metrics (table, below) show TnX has higher molecular complexity and better coverage uniformity compared to hyperactive Tn5. GC bias plots (right) also show improved coverage of in high and low GC regions. Normalized coverage over CDS exon regions (below) further show TnX has more even coverage compared to Tn5.

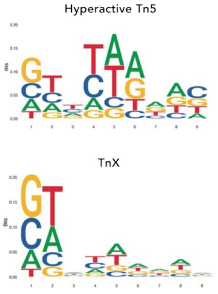
Enzyme	% Duplication	Estimated Library Size	CV of Coverage	Fold 80 Penalty	Fold 90 Penalty
Hyper Tn5	5.25%	1,345,701,144	36.1%	1.347	1.659
TnX	3.95%	1,992,181,229	32.5%	1.270	1.495





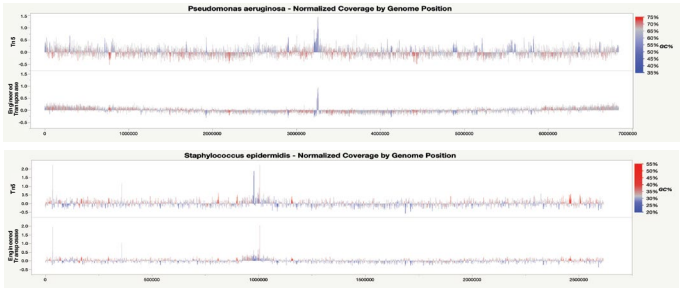
TnX has reduced Insertion Site Bias: Read start insertion bias was measured by examining the frequency of bases in the first 10 bases of each read. Logo plots per base position for hyperactive Tn5 and TnX illustrate reduced base bias in TnX (below, right). Insertion bias can be quantified by calculating the root mean square deviation (RMSD) at each position and plotting against the relative entropy for each, which is a measure of expected versus actual distribution of bases (below, left). Libraries made with TnX (blue) have lower overall base insertion bias than hyperactive Tn5 (red).





Performance on Difficult Bacterial Genomes

- Replicate bacterial WGS libraries were created from 100 ng of P. aeruginosa (69% GC) and S. epidermidis (32% GC) using library prep reagent sets made from hyperactive Tn5 and engineered transposases.
- Sequence coverage plots (below) for the two bacterial genomes demonstrate improved evenness of coverage in the engineered transposase libraries compared to Tn5 in both high and low GC regions.

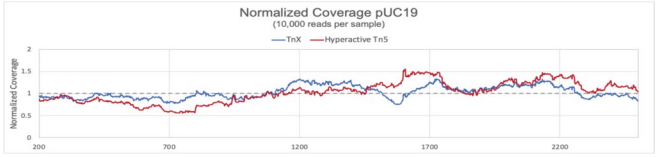


Performance in ExpressPlex™

- ExpressPlex is seqWell's rapid plasmid and amplicon library preparation workflow, formulated as a single "one pot" reaction that includes fragmentation, indexing, and amplification.
- An experimental set of ExpressPlex reagents were prepared using TnX and used to sequence pUC19 plasmids and compared to standard Tn5 ExpressPlex reagents.

Coverage of pUC19 using ExpressPlex made with TnX vs. Tn5 shows lower CV of coverage (table at right) and improved uniformity along the pUC19 sequence (bottom) for TnX.

coverage	TnX	Tn5 (control)
average	1036	1085
stdev	144	250
c.v.	14%	23%
max	1343	1643
min	750	587
Range (max / min)	1.8	2.8



Summary and Conclusions

- TnX transposase exhibits reduced insertion-site bias, increased activity and robustness compared to hyperactive Tn5.
- These features improve the data quality of high throughput transposase-based workflows by enabling better coverage uniformity, increased library complexity and the ability to formulate in stringent multi-reaction buffers for single reaction workflows.