Next Generation Synthetic Libraries for Enzyme Engineering, Cell Therapy and Gene Editing Technologies

Tavneet Gill, Xianan Liu, Tony Cox, Nathan Raynard, Caitlin Hoeber, Su Maw, Daniel Ramirez, Lucy Xu

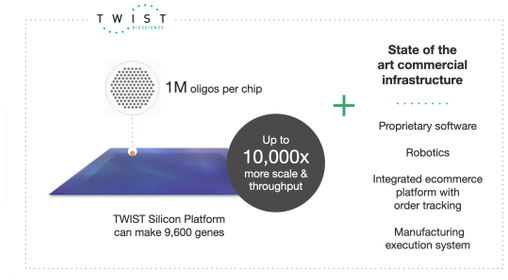
Abstract

Twist Bioscience has expanded a range of next generation library offerings which include CRISPR single guide RNA (sgRNA) and dual-guide (dgRNA) libraries, Synonymous Codon (SynCodon) libraries, and Paired T-cell receptor (TCR) and Chimeric Antigen Receptor T-cell (CAR-T) libraries. Utilizing Twist Bioscience's silicon DNA synthesis platform, precise mutations are individually synthesized and introduced via high quality oligo pools and gene fragments. Coupled with our molecular biology expertise in gene synthesis and DNA assembly, Twist Bioscience can build full-length, highly diverse, and precise libraries to maximize your screening capabilities.

CRISPR sgRNA and dgRNA, libraries are an efficient tool for high throughput gene editing and knockout of molecular targets. By leveraging the specificity of the CRISPR gene editing technology, researchers can combine their designed gRNAs and lentiviral plasmid, to create a ready-to-transduce library product. In recent literature, dgRNA libraries have become increasingly popular. The use of two gRNAs to target a single or multiple genes can help create a higher frequency and a higher specificity of mutations. With the use of efficient single stranded DNA isolation techniques, Twist can generate dgRNAs with >90% matching between two sets of gRNAs.

SynCodon libraries allow researchers to explore an emerging space in enzyme engineering. These libraries involve created mutations of a synonymous codon of each amino acid in a protein sequence. Codon optimization can help improve protein yields, binding affinity, stability, and expression of proteins. With the use of SynCodon Libraries customers can determine the most optimal synonymous codon usage for a given protein.

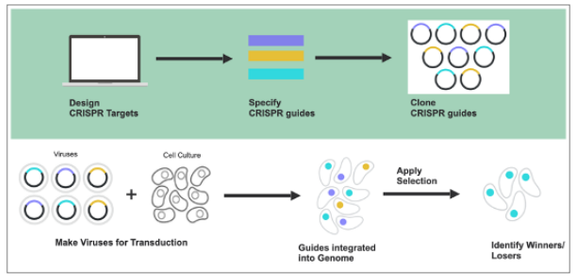
TCRs and Chimeric Antigen Receptor T-cell (CAR-T) libraries when assembled combinatorially can help screen a large scale of module sets within both library types. However with new long read sequencing technologies, customers have higher confidence in their alphas and beta pairing for TCR libraries and in conjunction antibody binding (scFv, Fab, VHH, scFab), hinge, transmembrane and signaling domain pairings for CAR-T libraries and therefore have specified designs for each module type. Twist utilizes a patented DNA assembly technology to effectively assemble unique combinations of domains in a high throughput approach to make downstream screening as precise and efficient as possible.



Synthetic libraries are based on computational in silico design and use synthetic DNA fragments to introduce diversity. Using Twist's unique silicon-based DNA synthesis platform, we can generate individualized modular domains for synthetic library production. These modules are constructed with novel DNA assembly technologies to provide diversities up to ~10^4 - 10^{10}

CRISPR gRNA Libraries

Twist Bioscience's CRISPR libraries allow for highly functional screening of customer defined guides that can come in the form of single or dual guide format. Twist leverages their internal customer vector onboarding process, to allow for custom lentiviral vectors to be integrated and used for current and future designs. With our high throughput silicon DNA synthesis platform we can synthesize gRNA libraries of up to 10E+06 complexity without compromising uniformity and representation. In addition, we offer cloning transformation efficiency of 100X of the desired diversity which can ensure variants are unbiased and abundantly represented in downstream screens. **Figure 1 A and Figure 1 B** summarize data collected from linear and cloned oligonucleotide pool of 1.52E+04 complexity. This GC plot shows the GC content of each variant (x-axis) and the normalized count for that particular variant (y-axis). From this graph, we can conclude that there is limited bias in variant integration based on GC content, and the variant pool is uniformly distributed in counts, with limited dropouts. In addition, we calculate the 95th/5th percentile, which shows us that 90% of the variants are within X multiplier of the mean. In figure 1A we see that the linear percentile metric is 1.99 which concludes that 90% of the variants are within 1.99X of the mean.



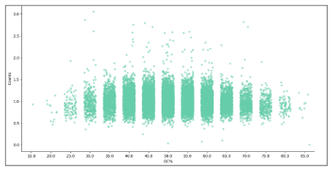


Figure 1A Linear: Percentile: 1.99 Dropouts: 1

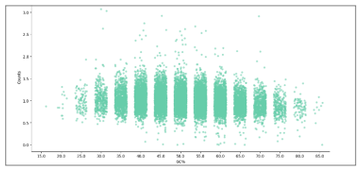


Figure 1B Cloned: Percentile: 2.08 Dropouts: 1

In addition to single guide libraries as the one represented in **Figure 1**, Twist also provides dual guide libraries to create higher specificity and efficiency in CRISPR screening technologies. One common problem with using single stranded oligo nucleotides to integrate perfect pairs of guides, is the potential chimeras that can form when amplifying out the oligonucleotide product. The nature of long oligonucleotide synthesis can generate a population of truncated material. In some cases "template switching" can occur, a phenomena where there are long conserved regions between oligonucleotide variants, that can potentially generate chimeras by introducing variable mis-priming events of single stranded DNA molecules. Twist has generated a process that can ensure chimera rates remain below 10% while maintaining high full length percentage of the DNA product. In **Figure 2** below we compare the chimera rate, on-target rate and truncation of a dgRNA library that contains +25,000 dgRNA pairs. The chimera rate is the percent of the pool that contains sequences where guide 1 does not match with its specified guide 2 but is in fact, full length. The on-target rate is the percent of expected dual guide pairs in full length. The truncation rate is the percent of reads that are not the expected length.
Ensuring that these dgRNA pools are highly uniform and perfectly paired will allow downstream screens to be much more efficient and specific.

Metric	Traditional Amplification	Twist In-House Amplification Method
Chimera Rate	38%	1.6%
On-Target Rate	53%	91.4%
Truncation Rate	10%	8%

Figure 2: Comparing an industry standard amplification format to Twist In-House amplification method. Here we can see that the chimera rate significantly improves in addition to the decreased rate of the truncation. Using this method the quality of and representation of the pool is preserved.

Synonymous Codon Libraries

Synonymous Codon (SynCodon) libraries provide a unique opportunity for protein exploration. By making specific and precise changes to a synonymous codon of a given amino acid, protein function can alter to have either increased binding affinity, stability, expression, or increased protein yields. Codon optimization is an efficient method to improve codon composition without creating mutations at the amino acid level. With tradition Twist Combinatorial Variant Libraries these changes be made in a concentrated domain or region within a specified protein. However with SynCodon libraries, changes can be made across the entire protein, allowing researchers to explore mutations along the entire protein that can influence downstream protein activity. SynCodon libraries are expression ready, and therefore top competitors can be screened phenotypically and targeted for single cell sequencing to genotypes of high performing codon compositions. In addition to fitness scoring, libraries can be measured using differential selection methods. The ideal workflow for SynCodon libraries would be to mutate codons in each position of the protein in Round 1 of library fabrication, then moved to targeting specific regions with revised codon frequencies for Round 2.

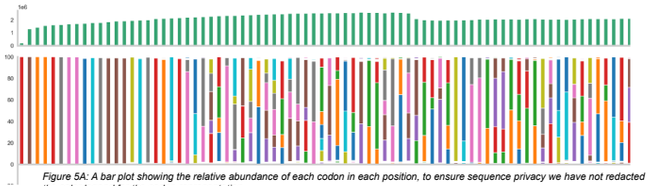


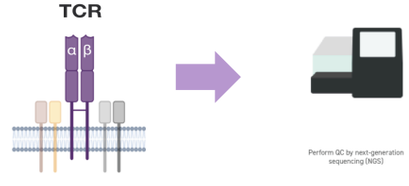
Figure 5A: A bar plot showing the relative abundance of each codon in each position, to ensure sequence privacy we have not redacted the color legend for the codon representation.

Figure 5B: A bar plot showing the representation of each oligo module to ensure they are uniformly represented and distributed and there are no biases between segments.

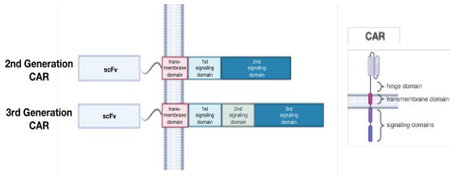
We perform quality control on SynCodon libraries to ensure that the mutations requested for the synonymous codons are represented, this is done through NGS sequencing coupled with a kmer analysis to resolve the DNA sequence across the entire library. In **Figure 5A**, we show the codon representation (y-xis) by position (x-axis). The bar plot at the top represents the coverage per position, so that we can ensure that the sequence is called with confidence and uniformity. In **Figure 5b** we show the representation of each designed oligo pool used in the fabrication process to ensure that there is a uniformity in the variants represented in the SynCodon library.

TCR and CAR-T Libraries

Engineered T Cell Receptor (TCR) and Chimeric Antigen Receptor (CAR) T cell therapies use the cells of our immune system for cancer therapeutics. Through the power of genetic engineering, these receptors can be modified, to provide a more personalized approach to cancer treatment. Emerging cellular sequencing technologies allow for researchers to identify sequences of target receptors and compare them to viable wildtypes. However, there are limitations to screening and testing thousands of sequences in downstream workflows. An alternative would be to adapt a modular approach to combinatorially assess multi-domain receptors. This allows researchers to test how the individual pieces work in tandem with one another to select the most optimal candidate for downstream drug development.
Twist Combinatorial Libraries present an agile approach to downstream screening workflows by allowing modular combinatorial assembly of TCR and CAR-T domains. Utilizing Twist Bioscience's silicon DNA synthesis platform coupled with the Combinatorial Variant Library assembly approaches, Twist can build full-length, highly diverse and precise libraries to maximize screening capabilities. These libraries are now offered in "perfect pair" format, where a corresponding beta chain can pair with appropriate alpha chain to limit the efforts needed for downstream screening.



Create paired or combinatorial libraries with your chosen gene inputs Customers with single-cell sequencing technologies/capabilities provide Twist team with TCR beta and alpha chain sequences. Twist will synthesize non-clonal fragments and assemble them with a constant linker using novel DNA assembly technologies. wist can help onboard a custom vectors into our internal database, that can be scaled up and conserved for future TCR library orders.



Find the most optimal combination of modular parts for applications like CAR Libraries Customers that would like to combinatorially assemble multiple modules can do so with Twist Bioscience. With the use of our DNA assembly technology, we can generate domains with strong uniformity, without compromising diversity. In addition, we can create CAR libraries that are non-homologous between individual module