Precision Synthesis of Variant Libraries Enables Comprehensive Interrogation of Single-Site Variant Space

Sonya Clark, Lucy Jia Xu, Helena Perazich, Shastyn Galaang, Weeney Chan
Twist Bioscience, 455 Mission Bay Blvd South, San Francisco, CA 94158

ABSTRACT

Single Site Variant Libraries (SSVL) enable the interrogation of sequence space to identify key residues in protein structure and function. Traditional methods of generating variant diversity, such as error- prone PCR and site-directed mutagenesis using degenerate or semi-degenerate oligonucleotides, generate only a random sampling of mutations in the gene sequence. These methods can create premature stop codons and incorporate redundant codons, which leads to poor control and uneven distribution of diversity. Twist Bioscience's silicon-based DNA synthesis platform, on the other hand, enables precise synthesis of variants to allow unprecedented, efficient, and comprehensive interrogation of the single variant space. Precision synthesis at the nucleotide and codon levels enables inclusion of desirable variants and exclusion of undesirable variants. Precision synthesis offers DNA sequence flexibility – it facilitates downstream manipulation, allowing restriction sites to be avoided. New data demonstrate the ability of Twist Bioscience's synthesis technology to construct nucleotide variant libraries and barcoded libraries for the in-depth examination of DNA or amino acid sequence space. The comprehensive, uniform incorporation of variants in each library is validated by next-generation sequencing, allowing informed conclusions to be drawn from both positive and negative experimental data. Combining the desirable diversity identified from Twist SSVLs in subsequent libraries provides an extremely efficient and intelligent way to navigate combinatorial sequence space.

WHAT CAN TWIST DO FOR YOU?

Decrease screening time and complexity with high-precision site variant libraries

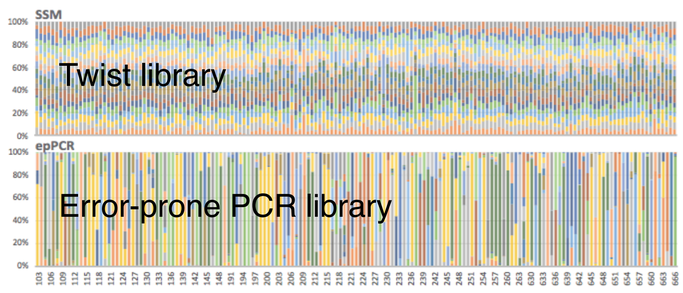
- Highly uniform representation due to precisely controlled introduction of each variant
- Minimal indel rates
- High on-target proportion
- Low off-target proportion (no stop codons due to high oligo synthesis fidelity)
- Complete codon selection control
- Complete avoidance of unwanted restriction sites and motifs
- Selective or saturation mutagenesis
- Full NGS verification of library composition

TYPES OF AMINO ACID AND NUCLEOTIDE SITE VARIANT LIBRARIES

Amino Acid Substitution

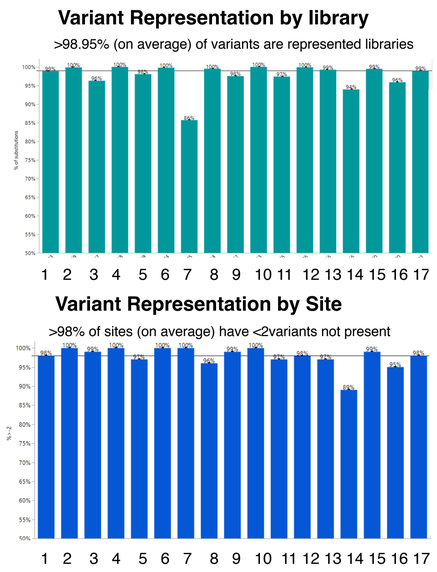
Application Spaces

- Enzymatic performance enhancement - improved specificity and fidelity, kinetics, substrate recognition, enantioselectivity
- Signal transduction - binding affinity modification
- Growth factor binding affinity optimization
- Antibody binding site modification
- Improvement in expression in difficult host types
- Protein robustness/stress tolerance



95% of variants are present between 2 and 8%

Twist provides an alternative strategy to interrogate protein sequence with uniform representation of each amino acid.



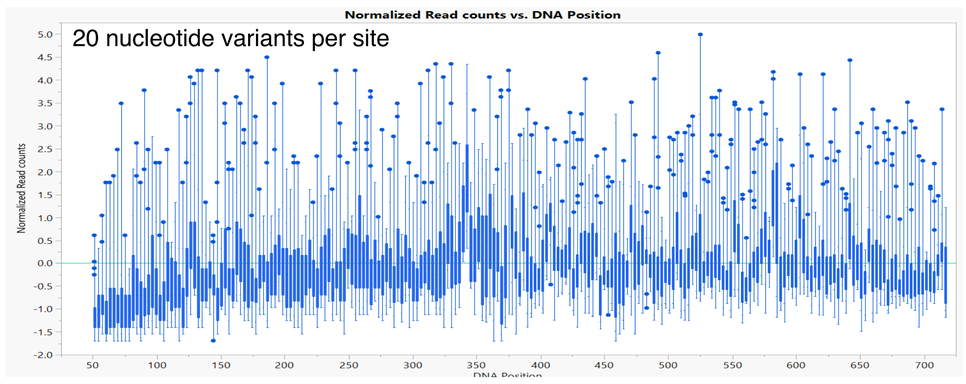
Advantages of TWIST amino acid substitution libraries

- Highly homogeneous representation of each introduced amino acid
  Traditional library generation methods (e.g., random mutagenesis and NNK/NNS libraries) have serious limitations, such as the introduction of bias, low control over diversity, and lack of precision
- High proportion of return of all desired variants
- No stop codons
- Complete control over codon usage at each position
- Codon choice (from all 64 possible) as per customer requirements
- Constructs from 350 bp to 4.5 kb (longer on request)
- 48 to >1,000 sites per library
- NGS sequence verification of requested variant representation

Nucleotide Substitution Libraries

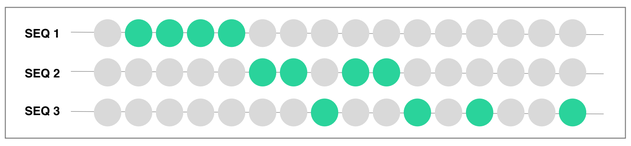
Application Spaces

- Promoter optimization
- Single base exploration for genetic element factor optimization
- Transcription modification
- Indel effects in non-coding regions
- Codon optimization



Advantages of TWIST nucleotide substitution libraries

- High and highly homogeneous representation of each introduced modification
- Modifications can be substitutions, insertions, or deletions
- Up to 40 modifications at each site
- Precision: no stop codons, no low-expression codons, low indels
- High proportion (>95%) of requested variants are present
- Codon choice is customizable
- NGS sequence verification

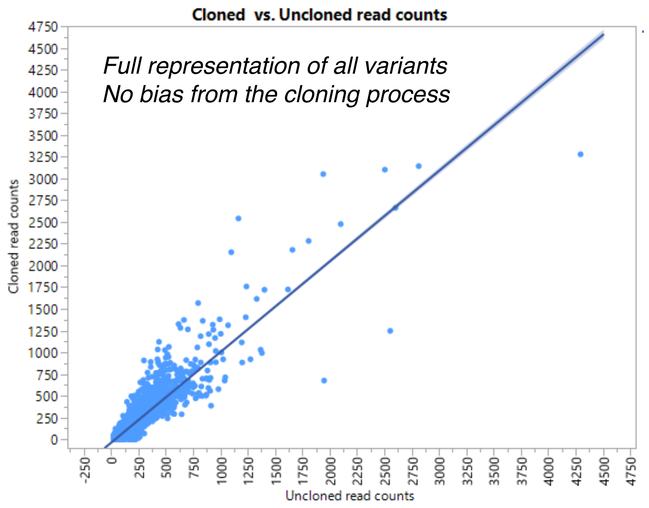


From 2 to 15 bases can be modified
(Either contiguous or non-contiguous changes within a 50 base region)

SITE VARIANT LIBRARY SERVICES

Cloning

- Delivered arrayed as one site/cloned per pool up to full library as one pool
- Fully validated non-biasing cloning process
- High quality full length clones at >90%
- Cloned into the vector of your choice or a Twist catalog vector
- Custom vector fully QC'd and integrated into Twists production system
- Go to cloned gene synthesis in one easy step!



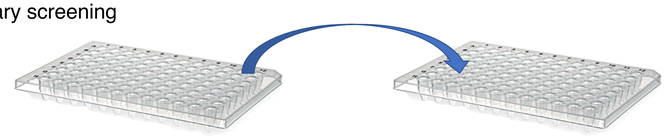
Next-Generation Sequencing (NGS) QC

- Detailed report of presence of requested variants
- >90% full length, indel-free
- All fragments verified at > 95% full length

Precision Barcoding (per Site)

Precisely **printed barcodes**

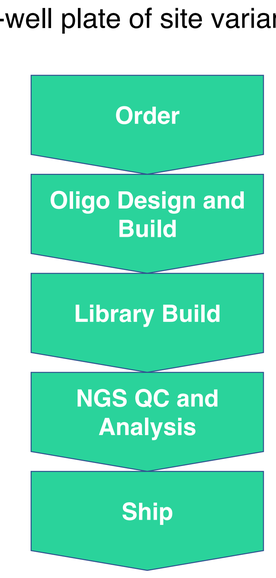
- Precision barcodes associate with variants at one barcode per site
- 5′ or 3′ barcode addition
- Low-cost option for decreasing complexity of library screening
- Sanger-validated barcode addition process



96-well plate of barcodes — 96-well plate of site variants

Customization

- Any number of sites
- Variable number of modifications per site (up to 20 amino acids, 40 codons)
- Codon usage
  - WT codon choice (original or synonymous)
  - Default = high expression of mammalian genes in *E. coli*
  - Top used codon applied in all cases
  - Defined codon choice (e.g., plant, bacterium of choice, etc.)
  - Restriction site avoidance
  - Design conflict identification (e.g., undesirable restriction site)
  - Flanking region addition to introduce unique termini
  - Restriction sites/cloning adaptor addition:
    - Max length 30 bases per side
    - Longer possible with extra charges
- Delivery format options: arrayed as single site per well or pooled



Twist Bioscience's massively parallel DNA synthesis platform, combined with advanced molecular biology technologies for library construction, creates more uniform and precise variant libraries than any other method for comprehensive generation and evaluation of functional and non-functional diversity for protein engineering.