# Site-Saturation Mutagenesis Library Construction Automated with the BioXp™ 3200 System

Steve Riedmuller, Senior Director Field Applications, Synthetic Genomics, Inc. La Jolla, CA

## Introduction

The BioXp™ 3200 System builds DNA fragments and plasmids hands-free from user sequence data. BioXp constructs can be used for a variety of downstream applications including multi-fragment assembly, gene construction and modification, and site-directed mutagenesis studies. Additional downstream applications of the BioXp system continue to emerge as new users discover and explore the instrument's potential. Here, we describe a novel application of the BioXp system— building a site-saturation variant library.

## Site-directed mutagenesis

First pioneered by Clyde A. Hutchinson III and Michael Smith[1], mutagenesis at specific DNA positions, also known as site-directed mutagenesis, has proven to be an invaluable research tool for elucidating protein structure and function. Site-saturation mutagenesis studies are exhaustive studies that sample the effects of randomly altering peptide sequences. Site-saturated mutagenesis has been applied to structure-function studies, directed evolution, and protein engineering, including substrate specificity analyses and enzyme activity alterations. A site-saturation variant library consists of a pool of constructs containing a random or targeted degenerate codon NNN across all or a set of sequences encoding the 20 standard amino acids. The variability can span a region of a protein of interest, or the mutations can be targeted at a particular amino acid site. The BioXp system has the capability to construct any of these targeted mutagenesis applications. Here we focus on usage spanning a region of a protein, altering single amino acids across a domain of a protein-of-interest.

## Traditional site-saturation library generation

Typically, degenerate primers are used to create site-saturation libraries with PCR amplification. Some drawbacks of these traditional library construction methods include the potential introduction of unintended errors by PCR, over-representation of amino acids encoded by redundant codons, the presence of premature termination codons, and the significant resources required for library screening. A major limitation of these types of studies is that the resources required to prepare and screen sites across a protein or region-of-interest can make the undertaking prohibitive, limiting usefulness of the technique for some research labs to one or very few sites at most, unless a library synthesis service is utilized. The benchtop BioXp instrument addresses some of these challenges, allowing users to generate vast NNN site-saturation libraries bench-side in a short period of time (less than 1 week for design, building DNA Tiles, cloning, and cell transformation) without the drawbacks of PCR.

## Building a site-saturation library on the BioXp system

One of the key advantages of automating DNA synthesis and construction is the increased capability to launch large-scale projects that may be too costly or labor-intensive to undertake using a service or with traditional methods. Taking full advantage of the capability of the BioXp system to build unique, high-quality linear DNA fragments (BioXp Tiles) in an overnight run allows users the opportunity to pursue more substantial studies. A case study demonstrating the large-scale capability and efficiency of the BioXp system is presented in this application note. For this study, a region within a protein-of-interest was systematically targeted. Specifically, we generated a site-saturation NNN library on a 432-bp construct across all 104 contiguous amino acids, targeting 1 amino acid position per reaction. The final library consisted of 104 individual pools, each representing randomization at a single unique amino acid position. A representation of a portion of this library is shown in Figure 1.

NNNAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACNNNGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATNNNGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTNNNCCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTANNNGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCANNNTCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCANNNGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTNNNCCCTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTNNNTCTACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCNNNACCAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTNNNAGCTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCNNNTGGAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCNNNAGAACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGNNNACTACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGANNNACTGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTNNNGCAACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTNNNACTAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCANNNAGA
AACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTNNN

**Figure 1. Example of a systematic site-saturation library targeting a single amino acid position in each library pool.** Each line of sequence represents an individual library pool constructed by the BioXp system. The targeted amino acid of each pool is designated by NNN, shown in red. For this study, 104 pools were generated across a 432 bp construct (19 "pools" of a representative 57-bp construct are shown).

## Methods

Following oligonucleotide design and pooling, reactions were initiated and run on the BioXp system. Representative results from gel electrophoresis of 2 sets of 48 consecutive pools are shown in Figure 2, demonstrating that the BioXp system provided consistent results across the sampling.

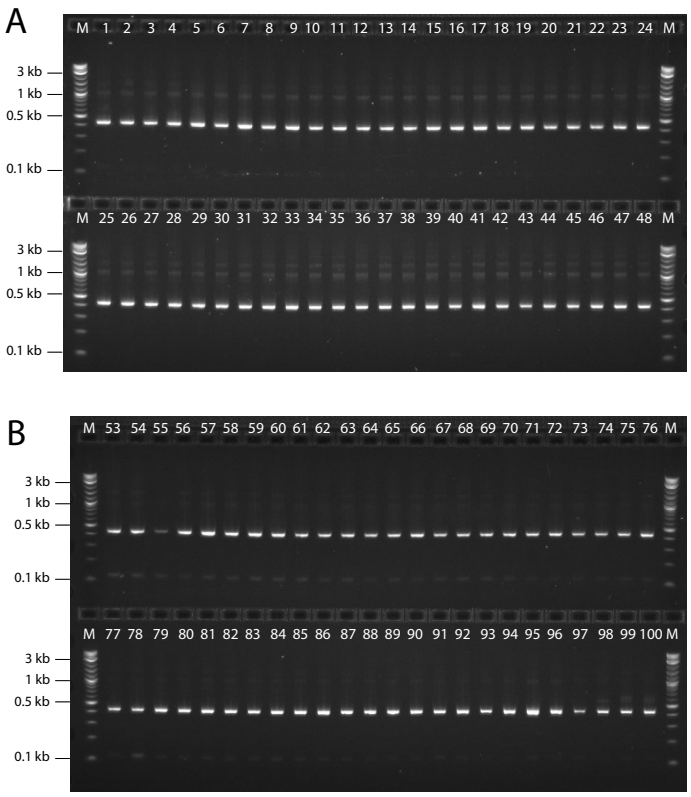


**Figure 2. The BioXp system provides consistent results across multiple runs and samples. (A)** Agarose gel electrophoresis of samples 1–48 (samples 49–52 exhibited similar results; data not shown). A prominent band at the expected 432 bp size is evident across all lanes. M designates the lanes in which 500 ng of 2-Log DNA Ladder (New England BioLabs) was loaded. **(B)** Agarose gel electrophoresis of samples 53–100 (samples 101–104 exhibited similar results; data not shown). A prominent band at the expected 432 bp size is evident across all lanes. Only sample 55 appears fainter than other samples, but this may be an artifact due to pipetting. M designates the lanes in which 500 ng of 2-Log DNA Ladder (New England BioLabs) was loaded.

## Results

To ascertain the fidelity of samples and randomization of mutagenesis, all 104 samples were pooled together for further analysis. From this library pool, 96 random constructs were Sanger-sequenced. As shown in Figure 3, we observed a high error-free rate of ~80% in the conserved regions and a sound distribution of NNN variation.

A

| Sequence Bin | Count of NNN | |
|---|---|---|
| 1-60 | | <--- Conserved priming region, no NNN expected |
| 61-120 | 21 | |
| 121-180 | 18 | |
| 181-240 | 24 | |
| 241-300 | 6 | |
| 301-360 | 18 | |
| 361-420 | 3 | |
| 421-480 | | <--- Conserved priming region, no NNN expected |
| Grand Total | 90 | |

B

| Sequence Bin | Count of Errors | |
|---|---|---|
| 1-60 | 1 | <--- Conserved priming region, reduction in errors expected |
| 61-120 | 11 | |
| 121-180 | 6 | |
| 181-240 | 5 | |
| 241-300 | 3 | |
| 301-360 | 6 | |
| 361-420 | 3 | |
| 421-480 | | <--- Conserved priming region, reduction in errors expected |
| Grand Total | 35 | |

**Figure 3. The BioXp system generates an NNN library with a high degree of variation and a high error-rate.** (A) The distribution of NNN variability from 96 random constructs is shown. (B) The frequency and distribution of errors are shown. These results show a high error-free rate of ~80% in the conserved regions.

## Summary

Generating a site-saturation NNN library on the BioXp instrument provides researchers a powerful new resource for high-efficiency, rapid library construction. For this study, a 432 bp construct that contains a 60 bp region of 26% GC was targeted for randomization. Our results demonstrate the generation of a library with substantial variability and a low error-rate. We view this study as a starting point for BioXp library construction and additional applications of the BioXp system. We anticipate that the BioXp system will enable additional expedited site-saturation library construction projects as well as other complex, library-based studies and analyses.

## Acknowledgements

We would like to thank Dr. John Gill for his assistance with design of the initial library and coordination of the post-transformation product analysis. We would also like to thank the BioXp scientific and engineering team for their help in the continued development of the BioXp instrument.

## Reference

1. Hutchison, C. A., Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P., & Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. Journal of Biological Chemistry, 253(18), 6551-6560.



SGI-DNA, a Synthetic Genomics Inc. Company
11099 North Torrey Pines Rd. Suite 150
La Jolla, CA 92037
Phone North America: 855-474-4362
Phone Outside North America: 858-228-4115