An Updated Duplex Sequencing Mutagenesis Assay Accurately and Reproducibly Detects Mutation Frequencies Below One-In-Ten-Million

IY Lee[1], DM Fitzgerald[1], JE Higgins[1]
[1]TwinStrand Biosciences, Inc., Seattle, WA.

TWINSTRAND® BIOSCIENCES

Mutagenesis Applications
(https://twinstrandbio.com)

Introduction

Duplex Sequencing (DS) is an advanced form of error-corrected Next-Generation Sequencing (ecNGS) technology that overcomes the limitations of traditional assays by accurately measuring and characterizing ultra-low mutation frequencies (MF) and mutation spectra in a wide range of in vivo and in vitro model systems. DS data have been shown to be concordant with transgenic rodent assays, and DuplexSeq™ Mutagenesis Assays are increasingly being adopted for preclinical mutagenicity assessment. Recent advances in DS technology have been formally incorporated into an improved assay, including DuplexSeq Mutagenesis Assay kits. In this study, we demonstrate that performance between previous (V1 + Enzymatic Fragmentation Chemistry) and improved assay (V2 Chemistry) designs is highly concordant, indicating the suitability of the updated DS assays for genotoxicity assessment.

Sequencing Errors Obscure Truth

Next-Generation Sequencing (NGS)

Single Strand Error-Corrected NGS

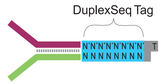
Duplex Sequencing

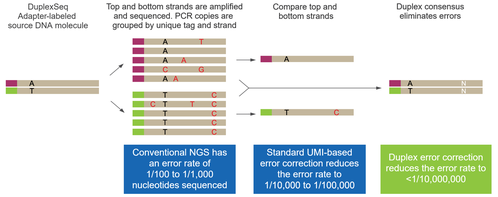
Duplex Sequencing eliminates errors to reveal ultra-rare variants that other sequencing methods miss.

TwinStrand Duplex Sequencing® Technology

A DuplexSeq Adapter has:

1. Identical (or relatable) degenerate tags in each strand.
2. An asymmetry allowing independent strand identification.

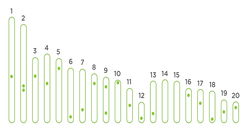




DuplexSeq Assay Versions:

- V1 + EF: the original chemistry formulation plus the enzymatic fragmentation (EF) module
- V2: formally incorporates EF into assay and streamlines workflow to be more user-friendly

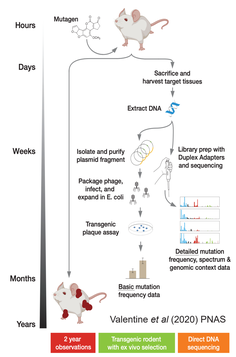
TwinStrand DuplexSeq Mutagenesis Assays



Mutagenesis Panels:

- 20 x 2.4 kb regions throughout genome
- Genic and intergenic regions included
- Regions not predicted to be under positive or negative selection
- Available for **rat** (shown on the left), **mouse**, and **human**
- Panel design is identical for V1 + EF and V2 chemistry assays

Measuring Genotoxicity



Tumor formation is triggered by exposure to genotoxic compounds, but tumors take years to form, making this an impractical end-point for routine safety testing. The ability of a compound to cause mutations can be assessed as a surrogate safety indicator. Transgenic rodent (TGR) assays are the current gold-standard for short-term *in vivo* mutagenicity testing, but require expensive animals and an extensive *ex vivo* selection process, and only yield basic mutation frequency data. **DS-based mutagenicity testing** can use wild-type rodents, such as those used in other steps of safety testing, and provides more comprehensive mutagenesis data.

Mutagens Used

BaP

- By-product of burning
- Requires metabolic activation
- Metabolized to epoxide which forms adducts at G

ENU

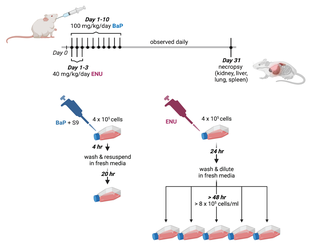
- "Model mutagen" and common positive control
- Direct alkylating agent

Abbreviations

DS = Duplex sequencing
MF = mutation frequency
BaP = benzo[a]pyrene
ENU = N-ethyl-N-nitrosourea

Part 1: Mutagenized Samples

Mutagen exposure protocols

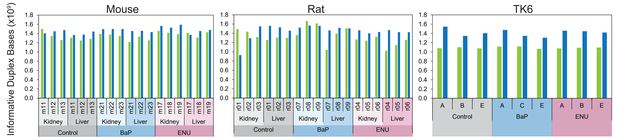


Male **Sprague Dawley rats** (n = 3 per group) and **C57BL/6Nac mice** (n = 3 per group) were administered **BaP** or **ENU** by oral gavage, according to the schedule illustrated (left, top). Untreated rodents were maintained as controls. Rodents were observed daily and kidney, and liver tissue were harvested at necropsy on day 31.

Human TK6 lymphoblastoid cells (n = 3 dishes per treatment) were exposed to **BaP + S9** or **ENU** as shown (left, bottom), or exposed to vehicle only. Doses were selected based on a range-finding study which targeted 30-60% cytotoxicity (CellTiter-Glo viability assay) and allowed for adequate growth during the post-exposure period.

DS was performed using 500 ng genomic DNA extracted from rodent tissues or cultured cells. Hybrid capture was performed with species-matched Mutagenesis Panels. Libraries were constructed using both V1+EF chemistry and V2 chemistry.

Consistent DS Assay Performance Metrics

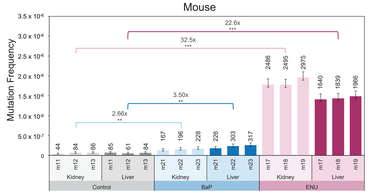


	Mouse		Rat		TK6 Cells	
	V1 + EF	V2	V1 + EF	V2	V1 + EF	V2
Avg Sequencing Clusters	163 M	165 M	169 M	171 M	156 M	158 M
Avg On-Target %	99.60%	99.40%	99.50%	99.30%	99.50%	99.50%
Avg Mean On-Target Depth	21,386x	23,081x	21,459x	23,026x	17,827x	22,693x
Avg Informative Duplex Bases	1.3 B	1.5 B	1.3 B	1.5 B	1.1 B	1.4 B

- Sequencing clusters: # of sequencing clusters assigned to each sample
- On-Target %: fraction of passing-filter bases located on or near the baited area
- Mean On-Target Depth: mean coverage of the target region
- Informative Duplex Bases: # of non-N consensus bases examined

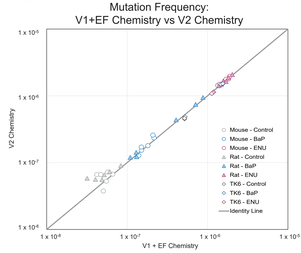
Baseline Mutation Frequencies < 1 x 10⁻⁷

$$MF_{min} = \frac{\text{number of unique variants}}{\text{informative duplex bases}}$$



Bars show per-sample mutation frequencies for each mouse library generated with V2 Chemistry, with Wilson binomial confidence intervals (95%). Numbers above bars represent total mutant bases. Numbers above brackets represent measured fold-changes of each tissue type per treatment group, compared to untreated control. MF for mouse kidney control samples were on average 5.7 x 10^{-8}, MF for mouse liver control samples were on average 6.4 x 10^{-8}. P-values were calculated from a quasi-Poisson generalized linear model comparing two groups, and adjusted for false discovery rate to account for multiple comparisons.
(**p value < 0.01, ***p value < 0.001)

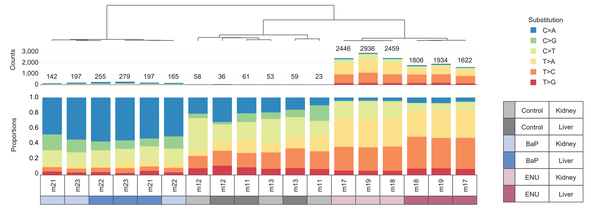
Highly Correlated MF Between V1+EF and V2 Chemistry



Scatter plot shows correlation of MF data generated by using V1 + EF and V2 chemistry from all samples. Mouse data are represented by circles, Rat data are represented by triangles, and TK6 cell data are represented by diamonds. Within each species, **control** samples are in gray, **BaP** samples are in blue, and **ENU** samples are in magenta.

V1 + EF and V2 chemistry data are highly correlated across different species and treatments. The coefficient of determination (R^2) was calculated for each species and is above 0.99 for each.

Single Base Substitution Spectra Perfectly Separated by Treatment



Unsupervised hierarchical clustering was performed within each species to group similar samples based on single base substitution spectra. All species' samples were perfectly separated by treatment group, and ENU samples are also perfectly separated by tissue (shown above is the mouse spectra generated from V2 Chemistry data). Numbers above the bars in the top panel are the number of unique single nucleotide variants detected in each sample.

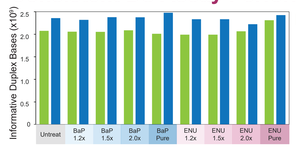
Part 2: Mixed Rat Samples

DNA mixture preparation

DS was performed on liver DNA from untreated rats and rats mutagenized with either **BaP** or **ENU** (1 animal each) to determine the MF of each pure sample. See "Part 1: Mutagenized Samples" for details on the unmixed DNA. DNA mixtures were made targeting mutation frequencies 1.2x (20%),1.5x (50%) and 2.0x (100%) higher than the untreated control, as illustrated (right). DS was performed on 1,000 ng of the DNA mixtures, in 4 replicates per mixture, using both V1+EF and V2 chemistry. Hybrid capture was performed with the rat Mutagenesis Panel.

Mixture calculations: $MF_{mixture} = (MF_{treated} - MF_{untreated}) * fraction_{treated} + MF_{untreated}$

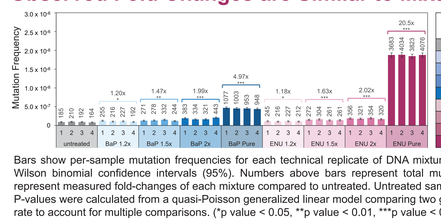
Consistent DS Assay Performance Metrics



	V1 + EF	V2
Avg Sequencing Clusters	264 M	268 M
Avg On-Target %	99.47%	99.09%
Avg Mean On-Target Depth	34,131x	39,009x
Avg Informative Duplex Bases	2.1 B	2.4 B

Bars show average Informative Duplex Bases across 4 technical replicates for each mixture using either version of the assay. See "Consistent DS Assay Performance Metrics" in Part 1 for descriptions of metrics.

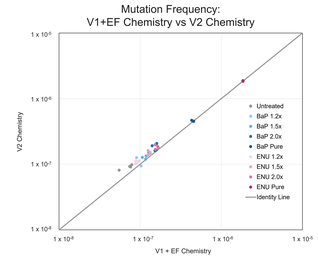
Observed Fold Changes are Similar to Mixture Targets



Sample Name	Target Fold Change	Observed Fold Change
Untreated	-	-
BaP 1.2x	1.20	1.20
BaP 1.5x	1.50	1.47
BaP 2.0x	2.00	1.99
BaP Pure	-	-
ENU 1.2x	1.20	1.18
ENU 1.5x	1.50	1.63
ENU 2.0x	2.00	2.02
ENU Pure	-	-

Bars show per-sample mutation frequencies for each technical replicate of DNA mixtures generated using V2 chemistry, with Wilson binomial confidence intervals (95%). Numbers above bars represent total mutant bases. Numbers above brackets represent measured fold-changes of each mixture compared to untreated. Untreated samples have an average MF of 9.1 x 10^{-8}. P-values were calculated from a quasi-Poisson generalized linear model comparing two groups, and adjusted for false discovery rate to account for multiple comparisons. (*p value < 0.05, **p value < 0.01, ***p value < 0.001)

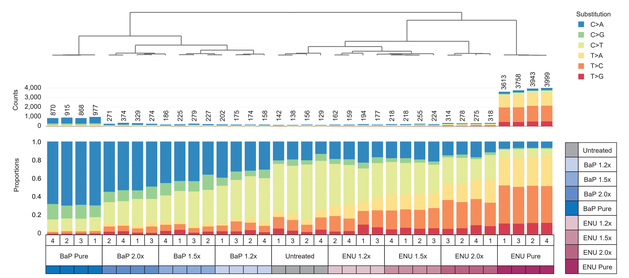
Highly Correlated MF Between V1+EF and V2 Chemistry Within Treatment Group



Scatter plot shows correlation of MF data generated by using V1+EF and V2 chemistry from DNA mixtures. **BaP** mixtures are in blue and **ENU** mixtures are in magenta.

V1+EF and V2 chemistry data are highly correlated across DNA mixtures. R^2 values for each treatment group are above 0.98 indicating high accuracy and linearity of mutation frequency measurements, even at relatively low mutation frequencies.

Single Base Substitution Spectra Perfectly Separated by Treatment and Target Fold Change



Unsupervised hierarchical clustering was performed to group similar samples based on single base substitution spectra. Each DNA mixture was perfectly separated and arranged in order of increasing target fold changes from the untreated control. Numbers above the bars in the top panel are the number of unique single nucleotide variants detected in each sample. This plot shows V2 Chemistry data.

Conclusions

- We quantify baseline MFs below 1-in-10-million in wild type rodents, which do not require complex transgenic breeding strategies.
- A significant increase in MF can be detected with n=3 animals per group.
- We demonstrate accurate and linear measurements of MF down to 1.2x induction, modeling detection of weak mutagenesis which can be seen in benchmark dose studies of mutagenic impurities.
- The updated assay generated very similar MFs compared to the previous version, while being more user-friendly.
- DS will continue to be valuable to researchers and regulatory agencies seeking to streamline and standardize genetic safety testing, reduce animal usage and cost, and accelerate time to insight.

Contact info:
https://twinstrandbio.com/contact/