Duplex sequencing for mutagenesis testing in wild-type rodents and a common human cell line

Devon M. Fitzgerald[1], Thao Huynh[1], Raul Burciaga[1], Camila Zanette[1], Jake Higgins[1], and Jesse Salk[1]
[1]TwinStrand Biosciences, Inc., Seattle, WA

TWINSTRAND BIOSCIENCES

Mutagenesis Applications
(https://twinstrandbio.com)

Introduction

Toxic compounds can alter the genetic makeup of organisms, generating mutations that cause cancer and other poor health outcomes. Standard mutagenesis assays yield a limited amount of indirect information about a compound's mutagenic potential and/or are complex and cumbersome.

Duplex sequencing (DS) is an error-corrected sequencing method that can directly detect, quantify, and characterize mutagenesis.

Here, we use TwinStrand DuplexSeq™ Mutagenesis Assays to characterize the mutagenicity of benzo[a]pyrene (B[a]P) and N-ethyl-N-nitrosourea (ENU) across multiple tissues of wild-type rodents and in a common cell line. Also, we generated abundant, characterized source material (genomic DNA) that can be shared with the community for inter-lab reproducibility studies.

This study paves the way for streamlined genotoxicology testing, reducing animal usage, cost, and time to insight.

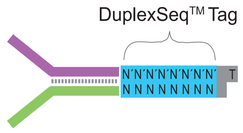
Sequencing Errors Obscure Truth

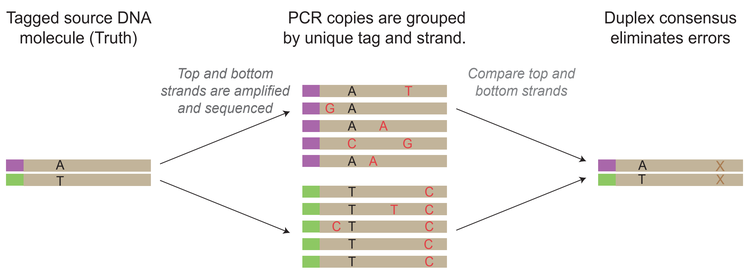
Next-Generation Sequencing (NGS) | Single Strand Error-Corrected NGS | **Duplex Sequencing**

TwinStrand Duplex Sequencing™ Technology

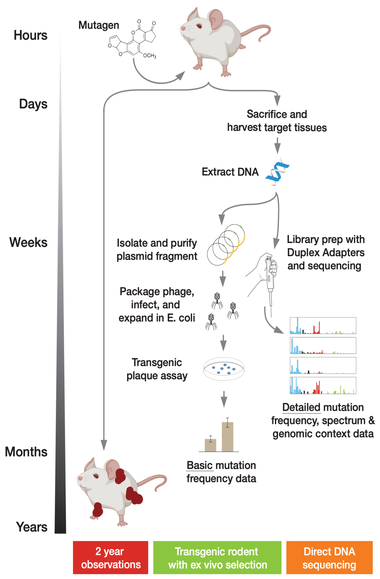
A DuplexSeq™ Adapter has:

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





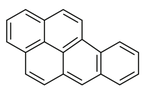
Measuring genotoxicity



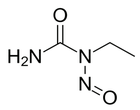
Valentine *et al* (2020) PNAS

Tumor formation is triggered by exposure to genotoxic compounds. But tumors take years to form, even in rodents, 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



B[a]P
- By-product of burning (*e.g.*, wood, cigarettes)
- Metabolized to epoxide which forms adducts at G

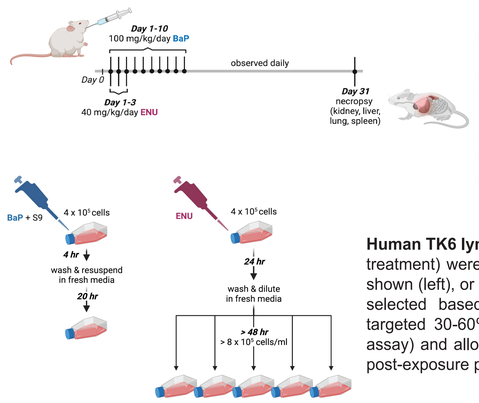


ENU
- "Model" mutagen used in screens
- Direct alkylating agent

Abbreviations

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

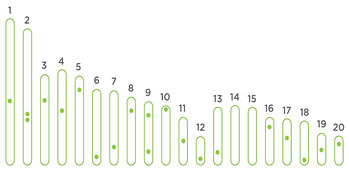
Mutagen exposure protocols



Male **Sprague Dawley rats** (n=3 per group) and **C57BL/6Ntac mice** (n=5 per group) were administered **B[a]P** or **ENU** by oral gavage, according to the schedule illustrated to the left. Untreated rodents were maintained as controls. Rodents were observed daily and kidney, liver, lung, and spleen tissue were harvested at necropsy on day 31.

Human TK6 lymphoblastoid cells (n= 5 dishes per treatment) were exposed to **B[a]P + S9** or **ENU** as shown (left), 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.

TwinStrand DuplexSeq™ Mutagenesis Assays

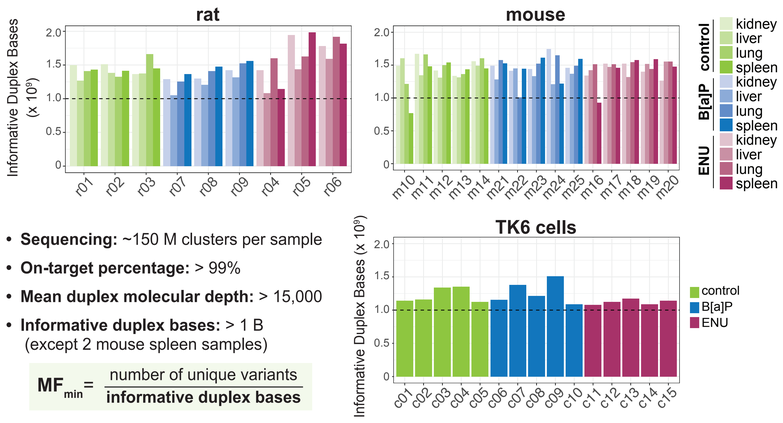


Mutagenesis Panels:

- 20 x 2.4 kb regions throughout genome
- genic and intergenic regions included
- regions not predicted to be involved in positive or negative selection
- available for **rat** (shown left), **mouse**, and **human**

Duplex Sequencing (DS) libraries were prepared using 500 ng genomic DNA extracted from rodent tissues or cultured cells. The Enzymatic Fragmentation Module (currently in early release) was used to construct all libraries, unless otherwise noted. Hybrid capture was performed with species-matched Mutagenesis Panels.

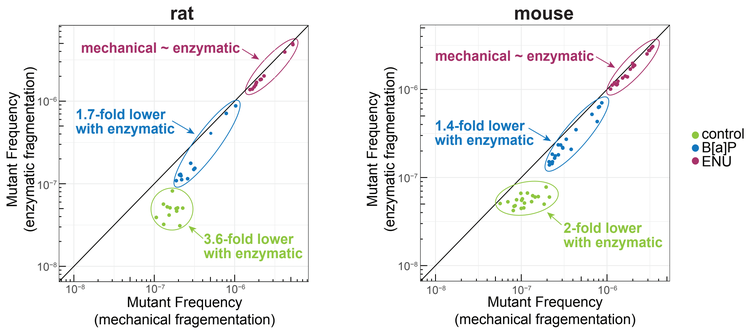
Duplex sequencing performance metrics



- **Sequencing:** ~150 M clusters per sample
- **On-target percentage:** > 99%
- **Mean duplex molecular depth:** > 15,000
- **Informative duplex bases:** > 1 B (except 2 mouse spleen samples)

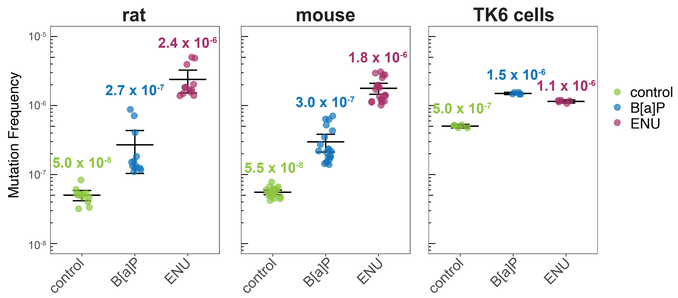
$$MF_{min} = \frac{\text{number of unique variants}}{\textbf{informative duplex bases}}$$

Enzymatic fragmentation "brightens" signal



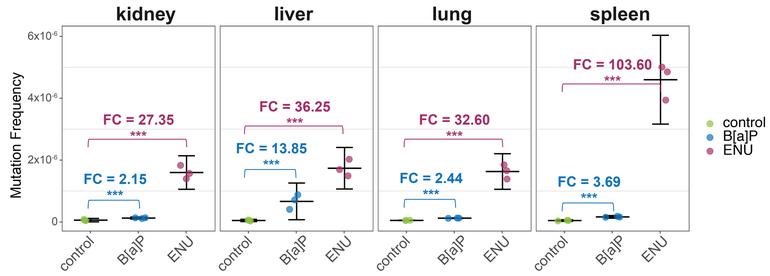
DS libraries were prepared using genomic DNA fragmented mechanically (Covaris) or with the new **Enzymatic Fragmentation Module** (early release). Enzymatic fragmentation reduces MFs of control samples, improving separation between control and treated groups and "brightening" the signal of mutagen exposure. Each data point represents a single tissue from an individual animal.

Mutation frequencies by species and treatment



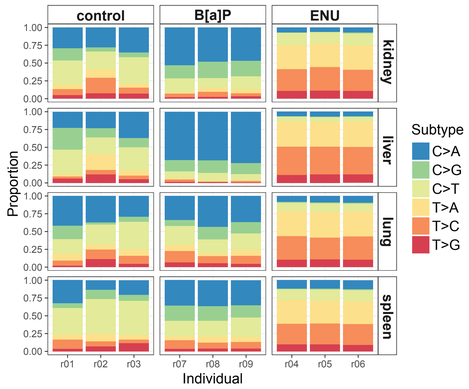
DS reveals clear differences in MF between treatment groups in all sample types. Absolute MFs are similar between rat and mouse. TK6 cells show a 10-fold higher basal MF than rodent tissues, but this does not obscure the mutagenic effects of **B[a]P** or **ENU**. Each data point represents a single tissue from an individual animal and values are plotted on a log scale. Error bars = 95% confidence intervals (CI).

Tissue-specific mutation frequencies (rat)



Both **B[a]P** and **ENU** treatments increase MF significantly across all rat tissues. The magnitude of mutagenic effects varies between tissues, with **B[a]P** causing the most mutagenesis in the liver and **ENU** having the most dramatic effect in spleen. *** $p < 2 \times 10^{-12}$, logistic regression; error bars = 95% CI.

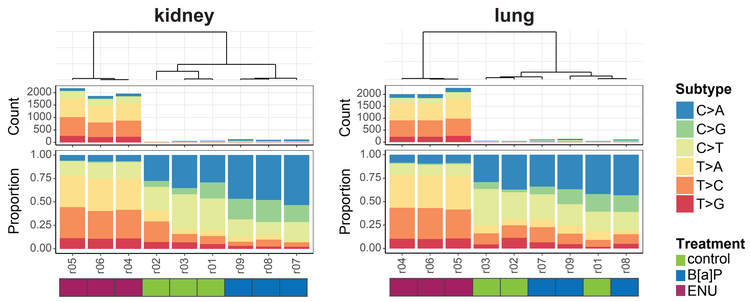
Single base-pair substitution spectra (rat)



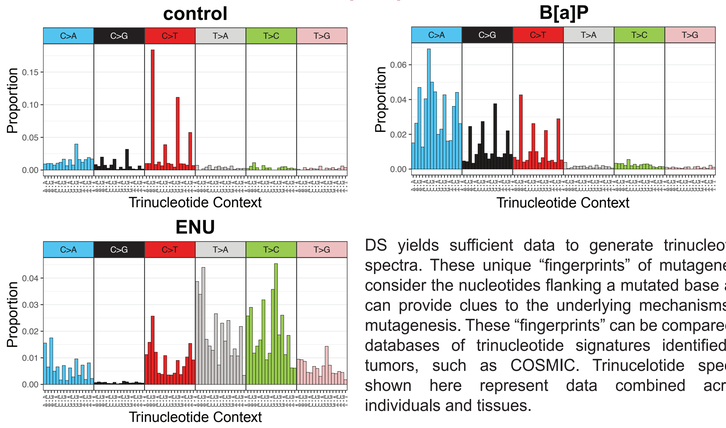
Grouping samples by treatment and tissue reveals differences in single base-pair substitution spectra (left). **B[a]P** treatment increases the proportion of C>A mutations. **ENU**-treated samples are dominated by T>A and T>C mutations. Spectra are consistent with previous reports.

Unsupervised clustering perfectly groups samples by treatment group in rat kidney (below left) and liver (not shown) and clearly separates **ENU** from other treatments in lung and spleen (below right and not shown).

Mouse and cell line datasets show similar single base-pair substitution spectra, with perfect clustering by treament (data not shown).



Trinucleotide spectra (rat)



DS yields sufficient data to generate trinucleotide spectra. These unique "fingerprints" of mutagenesis consider the nucleotides flanking a mutated base and can provide clues to the underlying mechanisms of mutagenesis. These "fingerprints" can be compared to databases of trinucleotide signatures identified in tumors, such as COSMIC. Trinucelotide spectra shown here represent data combined across individuals and tissues.

Conclusions

- DuplexSeq™ Mutagenesis Assays yield robust mutagenicity data from wild-type rodents and a common cell line after short-term exposure to two mutagens.
  - baseline mutation frequencies < 1-in-10-million among rodent controls
  - can detect subtle changes in mutation frequency (~2-fold change) and spectrum
  - single-base and trinucleotide spectra provide mechanistic insight into mutational processes
  - demonstrates potential to use same cohort of wild-type rodents for mutagenicity testing and other routine safety testing, reducing overall animal usage
- New Enzymatic Fragmentation Module boosts assay performance by "brightening" the signal of subtle mutagen exposure.
- Source material (genomic DNA) from this study is abundant and can be shared with the community for use in inter-lab validation studies.

Contact: https://twinstrandbio.com/contact/
dfitzgerald@twinstrandbio.com