DATA SHEET

TwinStrand DuplexSeq™ Mutagenesis Assays

Duplex Sequencing overcomes the limitations of conventional mutagenesis assays by providing users with novel insights into the genomic features and mechanisms that influence mutation induction and underlie the development of cancer.[1]

Product Highlights

- Directly detect and quantify DNA mutations using Next-Generation Sequencing (NGS)
- Compatible with *in vitro* and *in vivo* studies
- Three kits available to support human, mouse, or rat sample testing needs
- Standardized assay that provides the same readout regardless of sample type
- Accurate, reproducible, and highly sensitive measurement of mutations
- Includes cloud-based software for automated mutation analysis and report generation

TwinStrand DuplexSeq Mutagenesis Assays offer a fundamentally new approach for mutagenesis assessment via direct DNA sequencing and are compatible with any cell or tissue type across multiple relevant species. This approach enables quantitative and mechanistic mutagenesis testing using the same Next Generation Sequencing (NGS) instruments commonly used at the forefront of genomics and cancer research.

NGS is a powerful tool for detecting mutations in the genome. However, some types of mutations are only detectable by NGS when special error-correction approaches are used. Conventional NGS has an error rate of ~0.1–1% making it most appropriate for identifying clonal mutations that are present in a substantial proportion of the molecules being interrogated.[2] TwinStrand Duplex Sequencing™ technology is an extremely high accuracy form of error-corrected sequencing (ecNGS), which uses specialized adapters to independently track both strands of individual DNA molecules throughout library preparation and sequencing. The sequence reads derived from each original strand are compared to eliminate virtually all background errors. When applied to mutagenesis research and testing, duplex sequencing enables accurate and reproducible measurement of mutation loads on the order of 1–in–10 million.

An End-To-End Solution for Library Preparation and Analysis

DuplexSeq Mutagenesis Assays include all reagents necessary for DNA library preparation and target enrichment. DuplexSeq™ Software is hosted on DNAnexus®, a secure cloud-based platform, and includes an analysis pipeline that is optimized to provide a user-friendly data processing experience.

Figure 1. The streamlined DuplexSeq Mutagenesis Assay Workflow integrates with Illumina® instruments and workflows. DuplexSeq Software includes an analysis pipeline with automated report generation.

For Research Use Only. Not for use in diagnostic procedures.

© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Panel Design & Performance

DuplexSeq Mutagenesis Assays, available for human, mouse, or rat reference genomes, use a 48,000-base pair (48 kb) hybrid capture panel to target twenty 2,400 base pair (2.4 kb) regions. Each panel is designed to quantify somatic mutations, reflective of genome-wide mutagenic processes. The target regions are composed of an unbiased sampling of sequence contexts—GC%, trinucleotide composition, genic vs intergenic—which are representative of each respective genome. The regions have no known role in cancer and thus are unlikely to be significantly influenced by positive or negative selection. Repetitive elements are avoided to ensure high capture efficiency and overall sequencing quality. All probes and their respective target regions uniquely map to the reference genome. Technical optimization of the panels was performed to achieve a high on-target percentage and uniformity of capture.

Human Mutagenesis Panel: Location of the 20 Assay Target Sites in the Genome

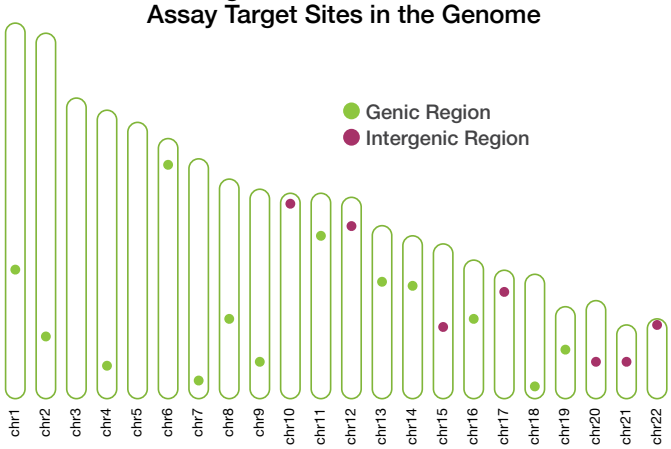

Figure 2. Location of genomic targets in the human mutagenesis panel. Mouse and rat panels are constructed similarly. Exact genomic targets can be found in the Panel Information Sheet for each assay.

Trinucleotide Composition of the Mutagenesis Panel and Genome are Proportional, Mitigating Sampling Bias

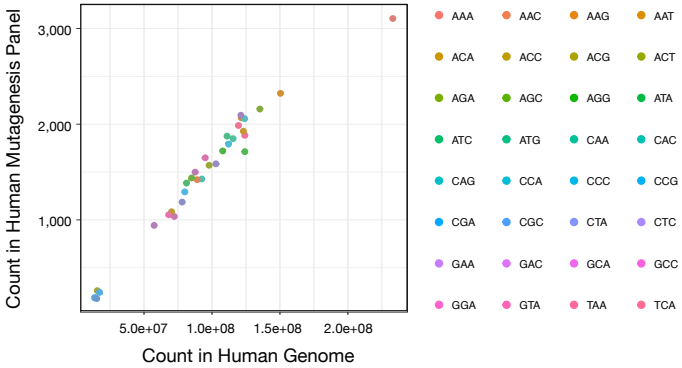

Figure 3. The count of each trinucleotide sequence in the human mutagenesis panel vs the human genome are shown at left. The composition of trinucleotides in the panel and genome are proportional, ensuring unbiased sampling when running DuplexSeq Mutagenesis Assays. Mouse and rat panels are constructed similarly.

DuplexSeq Mutagenesis Assays seamlessly integrate into existing NGS workflows without any major modifications or special equipment required. This standardized workflow is compatible with Illumina® NGS instruments. Library preparation, sequencing, and analysis can be completed in approximately one week.

Table 1: DuplexSeq Mutagenesis Assay Specifications

Parameter	Specification
Platform[a]	Illumina®
Target enrichment method	Hybrid-capture
Total probe footprint	48 kb
Number of target regions	20 (2.4 kb each)
Reference genomes	hs38DH (Human), mm10 (Mouse), rn6 (Rat)
DNA input range	250 ng – 2,000 ng
Total assay time[b]	3 days
Sequencing run[c]	2 x 150 cycles
Variant types detected	SNVs, MNVs, and indels
Uniformity of capture[d]	100%
Typical on-target %	95% - 99%

a. Due to sequencing depths required for DuplexSeq Assays, higher capacity instruments including NextSeq 550, NextSeq 2000, and NovaSeq 6000 are recommended. NovaSeq 6000 offers the greatest sequencing cost efficiency.

b. Three days account for library preparation only, including DNA shearing, two hybrid captures, and library quantification and pooling

c. 2 x 150 cycles are recommended for maximum depth and sensitivity.

d. Uniformity is defined as the percentage of bases in all targeted regions that are covered by at least 20% of the average base coverage depth reads.

For Research Use Only. Not for use in diagnostic procedures.
© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Interlaboratory Reproducibility

To evaluate interlaboratory reproducibility for the DuplexSeq Rat Mutagenesis Assay, fresh frozen rat liver samples were processed independently by TwinStrand and a commercial genetic toxicology reference lab study partner, MilliporeSigma. Big Blue® rats were treated with N-ethyl-N-nitrosourea (ENU), benzo[a]pyrene (B[a]P), or vehicle control (VC), following OECD Test Guideline No. 488 (N=6 rats per group).[3] DNA was extracted and split into two alliquots. Using the DuplexSeq Rat Mutagenesis Assay, a set of 18 samples were processed by each laboratory and sequenced. Mutation frequencies were assessed for each sample by both labs and compared. Near perfect correlation of mutant frequencies between sites was observed (R = 0.99) (Figure 4). Simple mutation spectra were also consistent between the two studies.

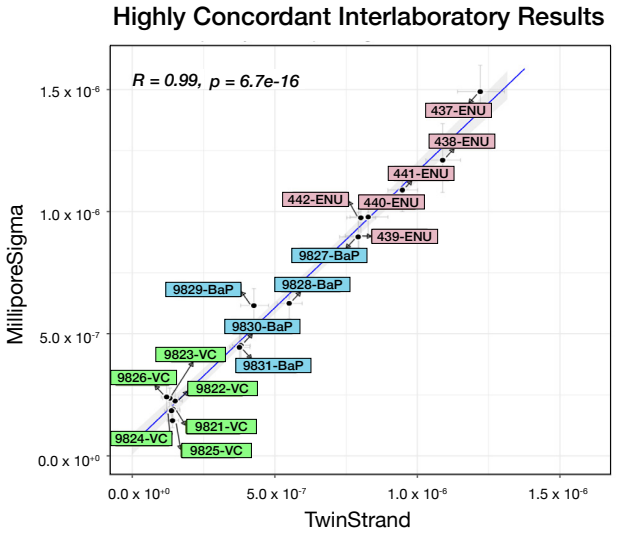

Figure 4. Mutant frequency correlation between study sites. DuplexSeq Rat Mutagenesis Assay performance is reproducible across independent labs.

Analysis & Reporting with DuplexSeq Software

Sequencing data are uploaded to the DNAnexus® platform for data processing. DuplexSeq Software includes an analysis pipeline that automatically generates mutation frequency, simple base substitution (SBS) spectra, and trinucleotide spectra plots in a comprehensive report. The trinucleotide spectra plots enable direct comparison of the mutation patterns observed in a sample to signatures reported in publicly available datasets, such as the Catalog of Somatic Mutations in Cancer (COSMIC) database, giving researchers a powerful tool to link mutagenic exposure to specific mutagenic mechanisms, and ultimately to human cancer processes.[4]

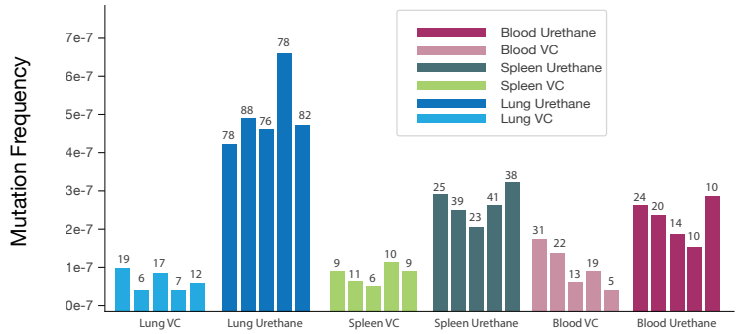

Figure 5. Duplex sequencing mutational analysis detected a significant increase in mutation frequency in tissue samples isolated from mice treated with urethane, a known mutagen, vs a vehicle control (VC). Mutation frequencies were highly reproducible across biological replicates (n=5/group).[5] The height of each bar indicates the sample mutation frequency (measured on the y-axis). The number of individual unique mutations observed per sample among the total number of DS BP sequenced is indicated above each bar.

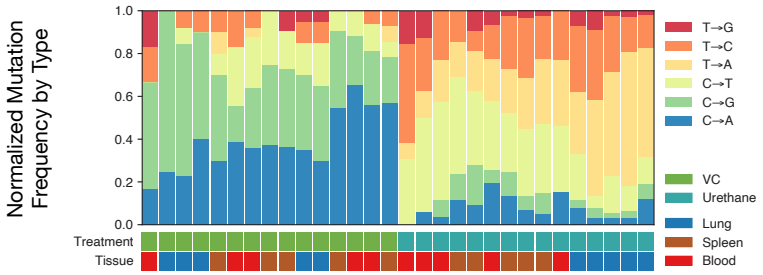

Figure 6. Simple base substitution (SBS) spectra provide a picture of the SBS patterns and identify similarities across samples. Mutation spectra were generated from duplex sequencing analysis of three tissues from mice exposed to urethane or VC. Unsupervised hierarchical clustering created groups that perfectly corresponded to the treatment groups (left, VC; right, urethane).[5]

Mutation Frequency Plots

Mutation frequency is a robust proportion metric used to normalize mutation data across samples. DuplexSeq Mutagenesis Assays enable the detection of many mutation types including single nucleotide variants (SNV), multi-nucleotide variants (MNV), and short insertions/deletions (indels). The mutation frequency of a sample is calculated by dividing the number of unique mutations observed by the total number of duplex base pairs sequenced (DS BP). The outputted mutation frequency plots show the calculated frequencies per sample.

Simple Base Substitution Spectra

The SBS spectra provide a measure of the mutation frequency of single base substitutions in a sample. DuplexSeq Software automatically groups spectrally similar samples using unsupervised hierarchical clustering. A color-coded mutational spectra plot (Figure 6), indicating the prevalence of specific base pair substitutions, is included in the report.

For Research Use Only. Not for use in diagnostic procedures.
© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Trinucleotide Spectra Plots

There are 96 different trinucleotide contexts in which a single base substitution mutation can occur.[7] Different mutagenic processes tend to generate more mutations in certain trinucleotide contexts and fewer mutations in others. By plotting the proportion of mutations observed across all 96 trinucleotide contexts, a pattern emerges which is characteristic of the mutagenic process(es) driving mutagenesis in the sample. DuplexSeq Software automatically generates a trinucleotide spectra plot (Figure 7) for each sample analyzed. The trinucleotide spectra can be assessed to identify mutational signatures in the samples, which are then compared to more than 60 unique trinucleotide signatures described in the COSMIC database to yield novel mechanistic insights into potential mutagenic mode of action.[4]

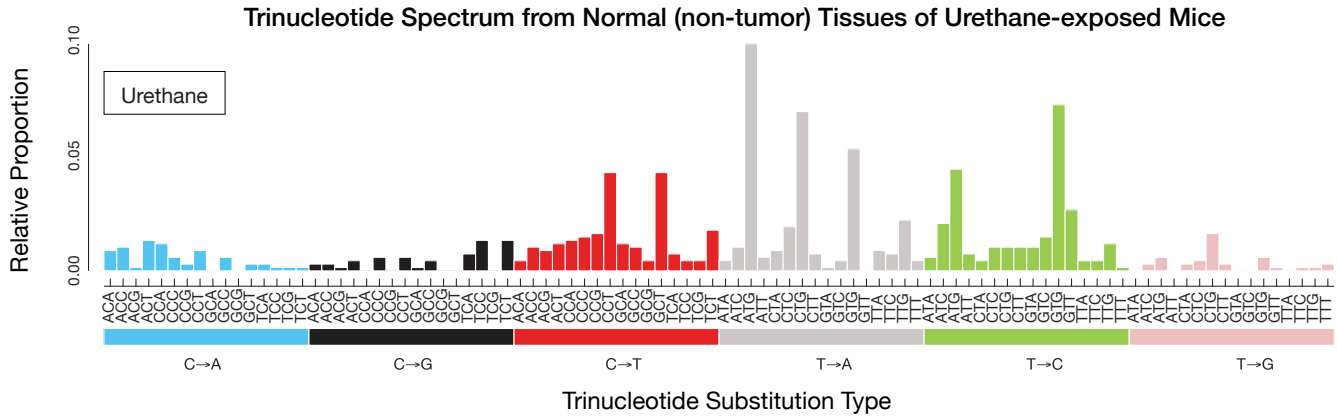

Figure 7. DuplexSeq was used to generate trinucleotide spectra from normal (non-tumor) tissues from urethane-exposed mice collected 28 days after exposure.[5] Underlying raw mutation data can be extracted from standard file outputs for statistical comparison to signature databases using publicly available relatedness calculators.

Sample Input Requirements

DuplexSeq Mutagenesis Assays are compatible with both *in vitro* studies involving specimens from cell lines, primary tissue culture or organotypic culture and *in vivo* studies, where fresh or frozen tissue from mouse, rat, or human specimens have been collected. Formalin-fixed, paraffin-embedded (FFPE) tissue not recommended for this high sensitivity application.

Generation of mutation frequency, SBS, and trinucleotide spectra plots require increasing amounts of sequencing data to generate plots with high confidence. In particular, the resolution of SBS and trinucleotide spectra plots increases with the number of mutations observed. The amount of inputted DNA can be increased to boost the amount of data generated and mutations observed. In general, for most mutagenized samples, a DNA input > 1,000 ng will be sufficient to generate a robust trinucleotide spectra plot.

Table 2: DuplexSeq Mutagenesis Assay Data Yield and Types of Data Plots Generated

DNA Input (ng)[a]	Data Yield[b]: Informative DS BP (Millions)			Types of data plots reliably generated with DNA inputs
	Human	Mouse	Rat	
250	310	350	320	Mutation Frequency
500	620	700	640	+SBS Spectra
1,000	1,240	1,400	1,280	+Trinucleotide Spectra
1,500	1,860	2,100	1,920	+Trinucleotide Spectra
2,000	2,480	2,800	2,560	+Trinucleotide Spectra

a. Assume a high-quality genomic DNA quantified prior to fragmentation. Data generated with mechanical DNA fragmentation.

b. Informative Duplex Sequencing Base Pairs (DS BP): the total number of non-N bases present among all duplex consensus sequences for a sample. This number serves as the denominator for calculating overall mutation frequency.

For Research Use Only. Not for use in diagnostic procedures.
© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Table 3: Estimated Number of Samples per Sequencer and Types of Data Plots Generated

DNA Input (ng)[a]	Sequencing Required, Clusters (Millions)[a]	Samples[b] / Sequencer[c] / Run[d]			Types of data plots reliably generated with DNA inputs
	Same for Human, Mouse, and Rat Panels	NextSeq™ 550	NextSeq™ 2000	NovaSeq™ 6000	
250	50	8	20	200	Mutation Frequency
500	100	4	10	100	+SBS Spectra
1,000	200	2	5	48	+Trinucleotide Spectra
1,500	300	1	3	32	+Trinucleotide Spectra
2,000	400	1	2	24	+Trinucleotide Spectra

a. One cluster equals two paired-end (PE) reads (e.g., 10 M clusters = 20 M PE reads). Data estimates assume the use of 150 bp PE read sequencer reagent kits.

b. Formalin-fixed, paraffin-embedded (FFPE) tissue is not appropriate for this high sensitivity application. Single-stranded DNA, nicks and DNA damage in FFPE risk generation of low-level pre-duplex artifacts as a result of the artificial conversion of ssDNA to dsDNA.

c. NextSeq 550 High Output (v2.5), NextSeq 2000 P3, and NovaSeq 6000 S4 (v1.5) are the instrument & flow cells represented. TwinStrand provides 48 unique dual indexes in assay kits. An additional set of 48 is available upon request.

d. Due to sequencing depths required for DuplexSeq Mutagenesis Assays, higher capacity instruments including NextSeq 550, NextSeq 2000, and NovaSeq 6000 are recommended. NovaSeq 6000 offers the greatest sequencing cost efficiency.

Data yield can be easily adjusted to meet experimental design requirements by modifying the quantity of DNA input. Using a higher input will yield more mutations and resolve spectral differences with higher resolution.

Advantages of DuplexSeq Assays

ecNGS approaches are rapidly emerging as powerful tools driving advancements in mutagenesis research and testing. DuplexSeq Mutagenesis Assays are the only commercially available NGS-based solution that quantifies non-clonal mutation loads accurately and reproducibly. The targeted panel design enables researchers to assess mutagenic effects throughout the genome.

Compatible with *in vitro* & *in vivo* study designs

Equally ideal for use with samples from *in vitro* or *in vivo* studies—including organoid and 3D tissue culture—DuplexSeq Mutagenesis Assays confer numerous advantages over conventional methods including improved sensitivity, accuracy, and variety of data generated—without the need for genetically modified organisms or cells.

Supports a range of genotoxicity testing needs

A standardized set of assays that provide the same readout across sample types to support genotoxicity testing needs from research and early safety screening through late development.

Provides insight into the mutagenic mechanism of action

The report includes the proportion of each base substitution type in a trinucleotide context, allowing the identification of unique mutational signatures.

Amenable to *in-human* use

DuplexSeq Mutagenesis Assays are the only commercially available solution compatible with all *in-human* tissues, including blood samples.

Integrated library preparation and data analysis workflow

DuplexSeq Assays seamlessly integrate with commonly used Illumina NGS workflows, and the sequencing data is then analyzed using TwinStrand's cloud-based data analysis software, hosted on the secure DNAnexus® platform.

For Research Use Only. Not for use in diagnostic procedures.

© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Key Publications: Mutagenesis

TwinStrand Duplex Sequencing™ Technology

Detection of DNA replication errors and 8-oxo-dGTP-mediated mutations in E. coli by Duplex DNA Sequencing

DNA Repair - 2023 | *Bhawsinghka et al.* | *DOI: 10.1016/j.dnarep.2023.103462*

DNA damage and somatic mutations in mammalian cells after irradiation with a nail polish dryer

Nature Communications - 2023 | *Zhivagui et al.* | *DOI: 10.1038/s41467-023-35876-8*

Error-corrected next-generation sequencing to advance nonclinical genotoxicity and carcinogenicity testing

Nature Reviews Drug Discovery - 2023 | *Marchetti et al.* | *DOI: 10.1038/d41573-023-00014-y*

Duplex sequencing identifes genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations

BMC Genomics - 2022 | *LeBlanc et al.* | *DOI: 10.1101/2022.10.21.513255*

"DS yields novel insights into the genomic features that influence mutation induction and consequently helps to elucidate potential mechanisms that underlie the development of human cancer. Furthermore, detailed spectra obtained from DS provides the opportunity to classify mutagen exposures based on the specific trinucleotide mutations induced. From a regulatory perspective, DS overcomes many of the limitations associated with conventional mutagenesis assays and can accurately and efficiently provide MF data to be used to inform sound regulatory decision making."

Genetic toxicity testing using human in vitro organotypic airway cultures: Assessing DNA damage with the CometChip and mutagenesis by Duplex Sequencing

Environ Mol Mutagen - 2021 | *Wang et al.* | *DOI: 10.1002/em.22444*

Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing

PNAS 2020 | *Valentine III et al.* | *DOI: 10.1073/pnas.201372417*

Mitochondrially-targeted APOBEC1 is a potent mtDNA mutator affecting mitochondrial function and organismal fitness in drosophila

Nature Communication - 2019 | *Andreazza et al.* | *DOI: 10.1038/s41467-019-10857-y*

Deleterious mitochondrial DNA point mutations are overrepresented in drosophila expressing a proofreading-defective DNA polymerase γ

PLoS Genetics - 2018 | *Samstag et al.* | *DOI: 10.1371/journal.pgen.1007805*

Mutational spectra of aflatoxin B1 in vivo establish biomarkers of exposure for human hepatocellular carcinoma

PNAS - 2017 | *Chawanthayatham et al.* | *DOI: 10.1073/pnas.1700759114*

Detection of ultra-rare mitochondrial mutations in breast stem cells by duplex sequencing

PLoS One - 2015 | *Ahn et al.* | *DOI: 10.1371/journal.pone.0136216*

Endogenous parkin preserves dopaminergic substantia nigral neurons following mitochondrial DNA mutagenic stress

Neuron - 2015 | *Pickrell et al.* | *DOI: 10.1016/j.neuron.2015.06.034*

Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage

PLoS Genetics - 2013 | *Kennedy et al.* | *DOI: 10.1371/journal.pgen.1003794*

For Research Use Only. Not for use in diagnostic procedures.

© 2023 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Ordering Information

Kit Type	Product Description	Number of Reactions	Product Number
Human	TwinStrand Duplex Sequencing™ Mutagenesis Panel (Human-50), v1.0	24	06-1005-02
		48	06-1005-03
Mouse	TwinStrand Duplex Sequencing™ Mutagenesis Panel (Mouse-50), v1.0	24	06-1006-02
		48	06-1006-03
Rat	TwinStrand Duplex Sequencing™ Mutagenesis Panel (Rat-50), v1.0	24	06-1007-02
		48	06-1007-03

To order your kit today or learn more about the TwinStrand DuplexSeq™ Mutagenesis Assays, contact us at www.twinstrandbio.com/contact.

References

1. LeBlanc DPM, et al. *BMC Genomics*. 2022 Jul;23(1):542
2. Salk JJ, et al. *Nat Rev Genet*. 2018 May;19(5):269-285.
3. OECD (2011), Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays, OECD Publishing, Paris.
4. Tate JG, et al. *Nucleic Acids Res*. 2019;47(D1):D941-D947.
5. Valentine CC 3rd, et al. *Proc Natl Acad Sci*. 2020;117(52):33414-33425.
6. Salk JJ, Kennedy SR. *Environ Mol Mutagen*. 2020;61(1):135-151.
7. Alexandrov LB, et al. *Nature*. 2020 Feb;578(7793):94-101.

TWINSTRAND™ BIOSCIENCES

3131 Elliott Ave, Suite 750 | Seattle, WA 98121 | +1(877) 202-TWIN | support@twinstrandbio.com

For Research Use Only. Not for use in diagnostic procedures. © 2023 TwinStrand Biosciences, Inc. All rights reserved.
All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

MKT0034 v1.0