SRSLY™ *cell-free* DNA NGS Library Preparation Kit for *Illumina*®

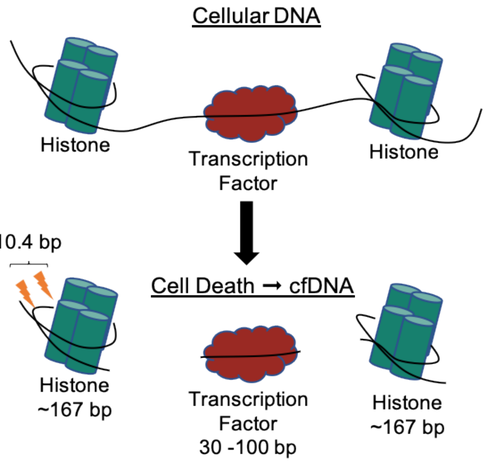
Kelly M. Harkins[1], Varsha Rao[1], Christopher J. Troll[1], Charles Cole[2], Colin Naughton[1], Jessica M. Morgan[1], Beth Shapiro[3], and Richard E. Green[2*]

[1]Claret Bioscience LLC, Santa Cruz, CA, 95060, USA, [2]Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA, [3]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA, *Corresponding Author

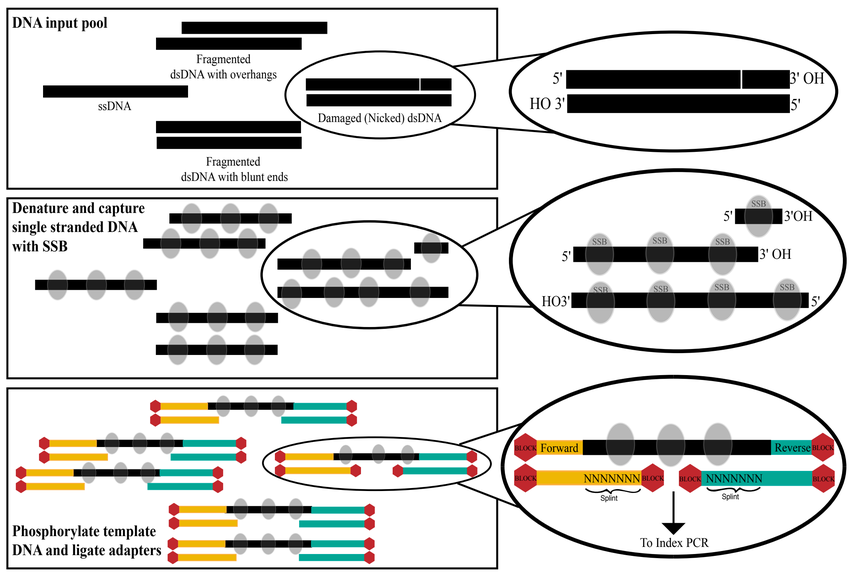
Introduction

Cell-free DNA (cfDNA) circulating in blood plasma and other bodily fluids **contains a wealth of biomedically relevant information** that can be assayed by next-generation sequencing (NGS) with a minimally-invasive blood draw. NGS data from cfDNA can reveal important aspects of cellular biology including prenatal health, organ transplant reception, and cancer detection and progression. However, **cfDNA is naturally fragmented, short and present in low abundance, creating obstacles for library preparation**, a requite step in the NGS workflow.

Single-stranded approaches to library preparation, initially developed for ancient DNA, **capture higher proportions of short and degraded DNA fragments compared to traditional double-stranded methods**. This feature makes single-stranded approaches ideally suited for cfDNA NGS applications. However, widespread adoption in the NGS community has been hindered because single-stranded methods are more time consuming than double-stranded methods, require exotic or single-source reagents, and in some cases generate artifacts that require downstream data processing. Until now…



The SRSLY Workflow



HERE WE PRESENT SRSLY, a simple and efficient ligation-based ssDNA library preparation method that is engineered to produce complex libraries from low inputs of cfDNA without alteration to the native ends of template molecules. SRSLY works in a one-step combined phosphorylation/ligation step that simultaneously prepares template DNA molecules for ligation without end-polishing and ligates proprietary splint adapters compatible with Illumina platforms.

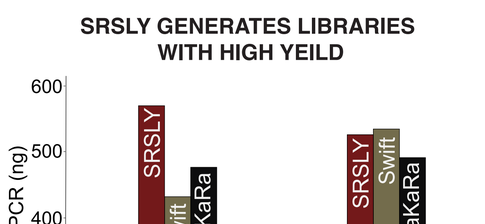
SRSLY Features

- **Recovers duplex DNA as well as single- stranded and nicked dsDNA**
- **Optimized for 1ng of cell-free DNA, with low end inputs as low as 50pg/uL**
- **Single reaction reduces errors, bench time and is amenable to automation**
- **No end-polishing preserves natural DNA fragment ends**
- **Superior recovery of short fragments**
- **Form DNA to sequence ready Illumina libraries in under 3 hours**
- **Compatible with single and dual indexing, as well as unique molecular identifier (UMI) incorporation**

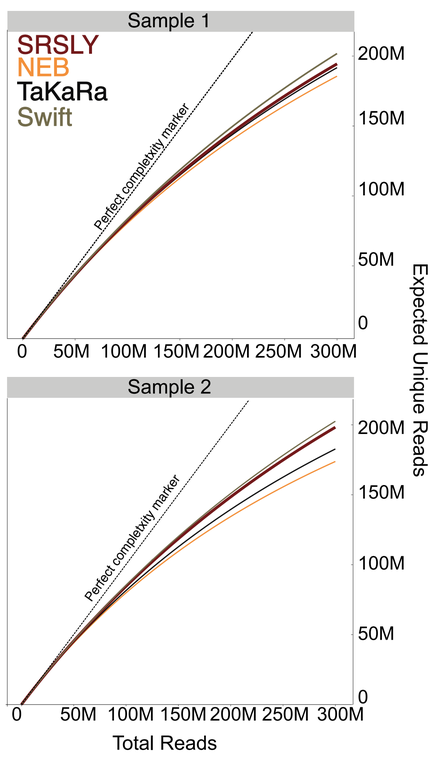
Downstream Applications
- Exome Sequencing
- Panel Enrichment
- Nucleosome Positioning
- SNP Calling
- Novel Discovery

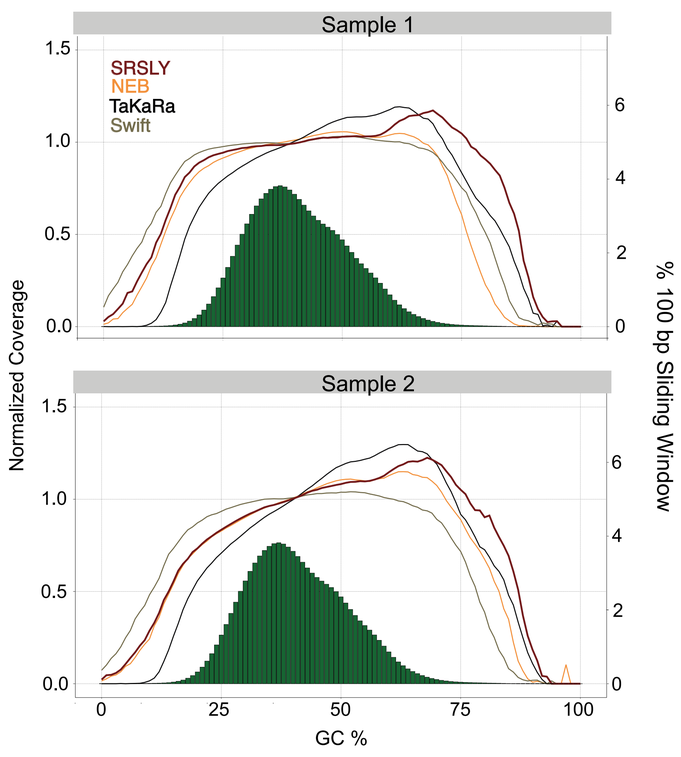
Fields of Use
- Liquid Biopsy
- Oncology
- Prenatal Testing
- Transplant Medicine

SRSLY NGS Metrics Rivals or Outperforms Other Commercial Kits



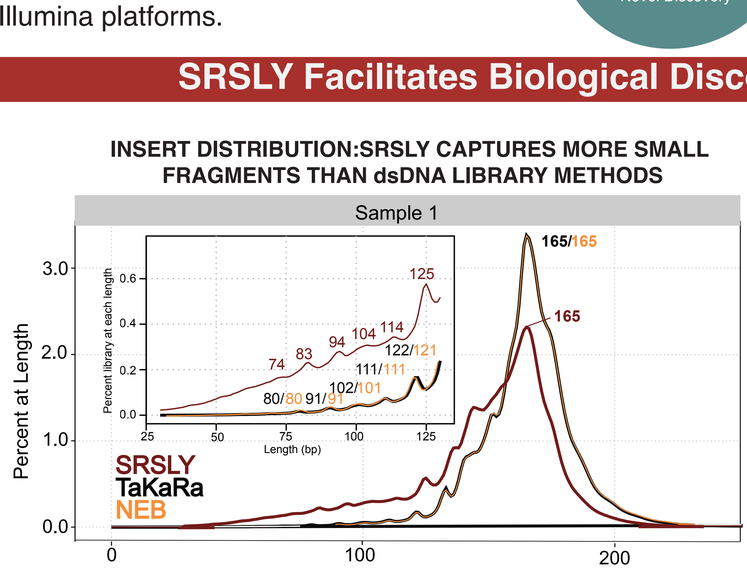
Kit	Type of prep method	Input cfDNA amount	Total time through index PCR	Input cfDNA sample ID	Yield post 10c index PCR (Total)	Merged read pairs sequenced	Mapping rate	Mapped reads in 30-100bp bin
ClaretBio SRSLY™	Single-stranded	1ng	2.5hrs	Sample 1	569ng	73,077,812	7.77%	94.35%
				Sample 2	525ng	63,449,167	12.0%	93.52%
NEBNext® Ultra II™	Double-stranded	1ng	3hrs	Sample 1	278ng	78,705,033	0.49%	94.66%
				Sample 2	310ng	54,734,560	0.49%	88.88%
TaKaRa ThruPLEX® Plasma-Seq	Double-stranded	1ng	2.5hrs	Sample 1	476ng	74,429,274	0.77%	94.01%
				Sample 2	470ng	71,294,525	1.32%	87.97%
Swift Accel NGS® 1S Plus	Single-stranded	1ng	3.5hrs	Sample 1	432ng	78,018,897	6.94%	93.66%
				Sample 2	534ng	69,959,246	10.68%	93.73%

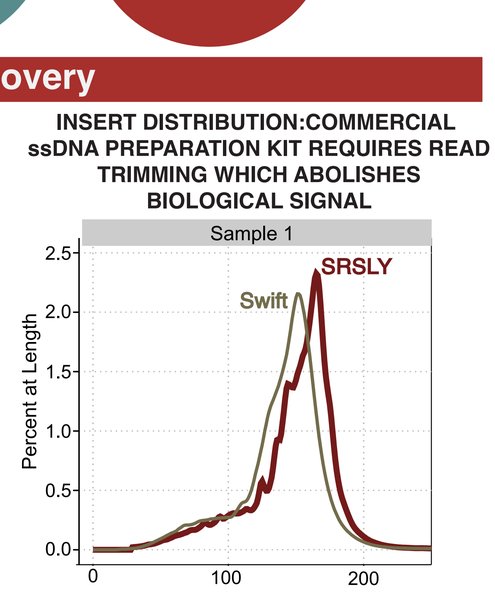


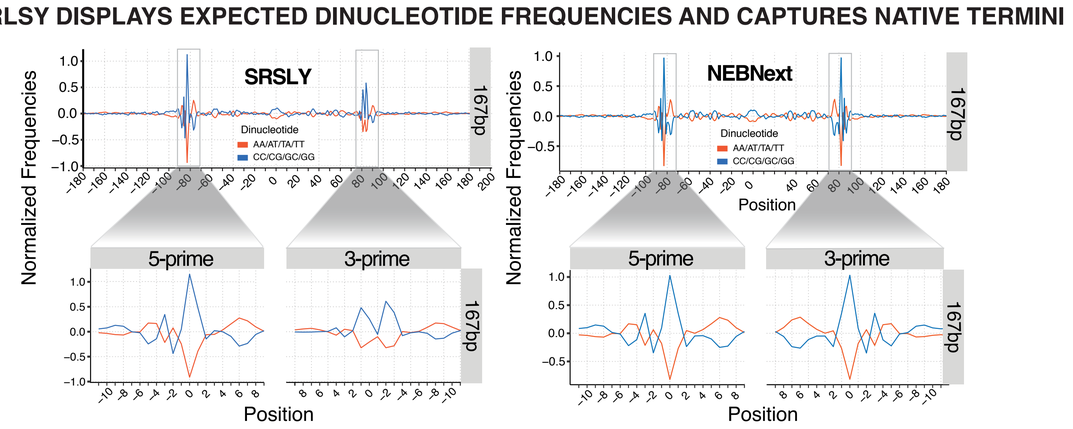


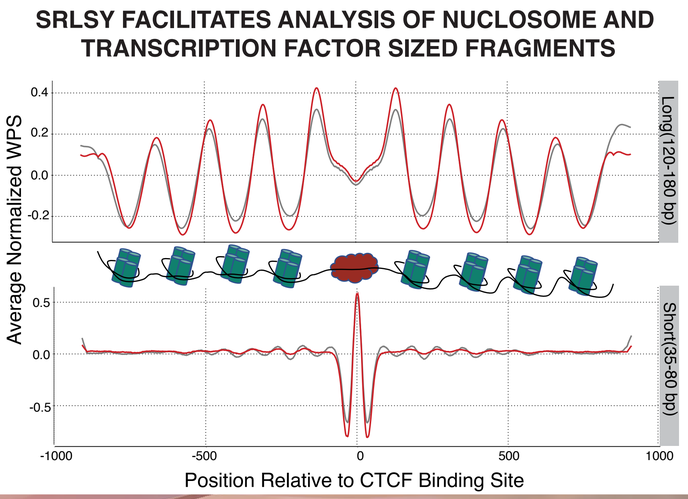
cfDNA libraries were prepared using SRSLY and commercial kits. **Yeild.** Libraries quantified using a Qubit 3.0 with 2 µl of final purified library post 10c index PCR. **Mapping.** Libraries sequenced 2x151 on a HiSeq X. Reads were adapter trimmed and merged (SeqPrep 2.0) and mapped (bwa aln) to hg19 q > 20. **Small Fragments.** Binned according to size (30-100bp; >100 bp). **Complexity.** (PreSeq lc_extrap) was estimated; the number of unique reads was extrapolated to 300M read-pairs, roughly the yield from one lane of HiSeq X. **Normalized GC coverage**. calculated using Picard Tools CollectGcBiasMetrics.

SRSLY Facilitates Biological Discovery









Insert distributions. Insert lengths are calculated from merged, mapped reads (q > 20).**Dinucleotide counts.** Fragmentation points were calculated (custom python script) for either a 100 bp or 11 bp window, where 100 bp or 11 bp of genomic context at both 5-prime and 3-prime fragmentation points were added respectively. The data was normalized using a median filter and dinucleotide frequency was plotted for weak (AA/AT/TA/TT) vs strong (CC/CG/GC/GG) dinucleotide interaction such that the center of the insert was at 0 and the regions upstream and downstream of the fragmentation point had negative and positive values, respectively. **CTCF binding sites.** Window Protection Scores were calculated in the manner described in *Snyder et al, 2015*. A bed file containing a list of putative TF binding sites was downloaded from the JASPAR2018 table (hub_186875_JasparTFBS) from the UCSC Genome Browser Table Browser and filtered to include only CTCF sites used by *Synder et al, 2015* (PMID: 26771485).

Interested In Our cfDNA Beta Program?

Visit: www.claretbio.com/contact

Scan here for our SRSLY bioRxive publication

Scan here to visit our website