illumina®

TruSeq® ChIP Library Preparation Kit

Proven TruSeq data quality delivers the most complete and accurate profile of target protein–DNA interactions.

Highlights

- **Proven TruSeq Data Quality**
  Most complete and accurate profile of target protein: DNA interactions
- **Low DNA Input Requirement**
  Robust results from just 5 ng DNA from a range of sample sources
- **Simple, Streamlined Workflow**
  Enhanced scalability with an easy-to-use, simplified workflow
- **Multiplexed Sequencing with 24 Available Indexes**
  Optimize sequencing output distribution across samples, reducing cost per sample

Introduction

Determining how protein–DNA interactions regulate gene expression is essential for fully understanding many biological processes and disease states. This epigenetic information is complementary to DNA sequencing, genotyping, gene expression, and other forms of genomic analysis. Chromatin immunoprecipitation sequencing (ChIP-Seq) leverages next-generation sequencing (NGS) to quickly and efficiently determine the distribution and abundance of DNA-bound protein targets of interest across the genome. ChIP-Seq has become one of the most widely applied NGS-based applications, enabling researchers to reliably identify binding sites of a broad range of targets across the entire genome with high resolution and without constraints.

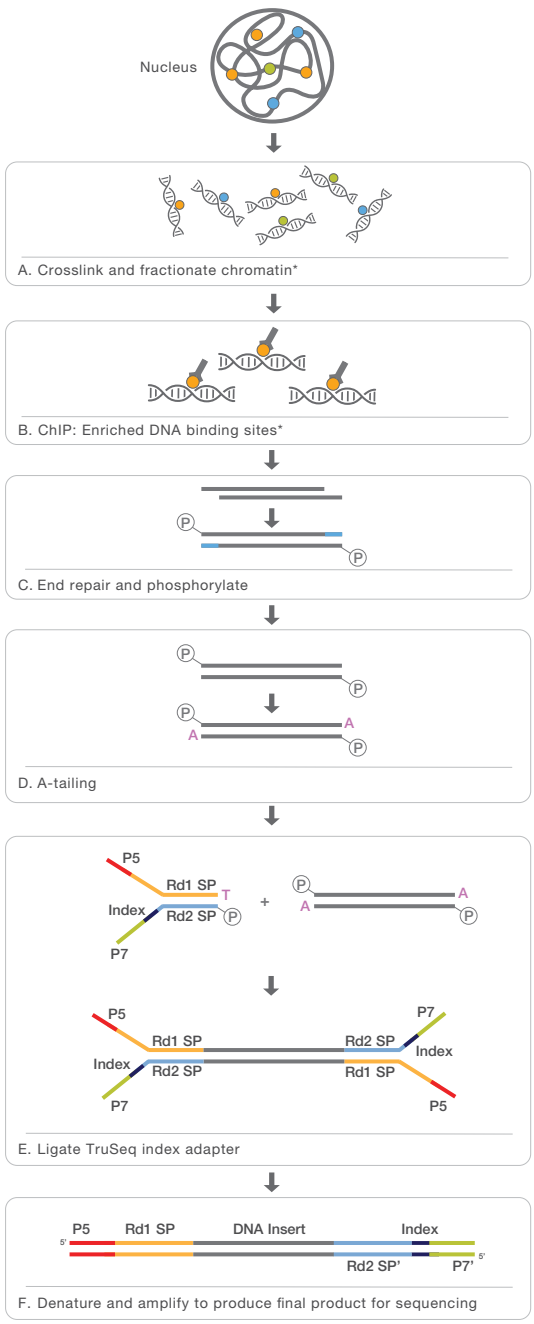
As the output of NGS systems has increased, ChIP-Seq researchers increasingly require a combination of highly multiplexed sequencing and simple, streamlined workflows. TruSeq ChIP Library Preparation Kits meet those demands, offering a simple, cost-effective solution for obtaining visibility into the mechanics of gene regulation. Library generation from ChIP-derived DNA includes the addition of indexed adapters, enabling the optimal distribution of sequencing output based on coverage needs. An optimized, highly scalable library preparation workflow and master-mixed reagents reduce hands-on time and support an automation-friendly format for parallel processing of up to 48 samples. Samples with different indices can be mixed and matched to maximize experimental throughput. A low sample input requirement (5 ng) ensures robust results even when input DNA availability is limited, providing flexibility in the choice of sample source and target proteins for analysis.

Simple, Streamlined Workflow

TruSeq ChIP Library Preparation Kits provide a significantly improved library preparation workflow compared to other methods. The TruSeq workflow reduces the number of purification, sample transfer, pipetting, and clean-up steps. A universal adapter design incorporates an index sequence at the initial ligation step for improved workflow efficiency and more robust multiplex sequencing (Figure 1).

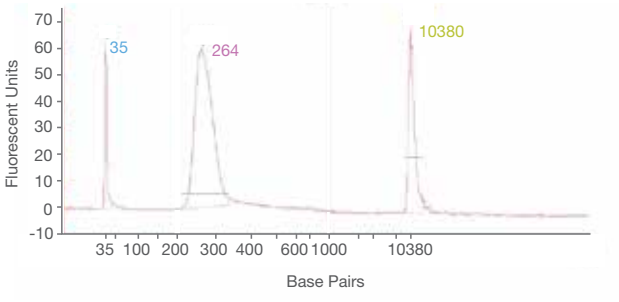
*Steps A and B are performed prior to the TruSeq ChIP Library Prep workflow.

Figure 1: ChIP-Seq Workflow



The simple, streamlined TruSeq ChIP Library Preparation Kit workflow (Steps C–F), reduces hands-on time and speeds analysis.TruSeq universal adapters improve workflow efficiency and enable robust multiplex sequencing.

Figure 2: Bioanalyzer Trace of MafK Library



Bioanalyzer trace data for a library generated for transcription factor target MafK using the TruSeq ChIP Library Preparation Kit with 5 ng of input DNA. The center peak indicates robust yield within the desired insert size range.

The TruSeq ChIP process begins with the enrichment of specific cross-linked DNA-protein complexes using an antibody against a protein of interest (Figure 1A-B). The stretches of DNA bound to the target protein are then isolated and used as input DNA for library generation. DNA fragments are end-repaired and an 'A'-base added to the blunt ends of each strand, preparing them for ligation to the sequencing adapters (Figure 1C-D). Each TruSeq adapter contains a 'T'-base overhang on the 3'-end providing a complementary overhang for ligating the adapter to the A-tailed fragmented DNA (Figure 1E). Final product is created (Figure 1F) and after size selection, all of the ChIP DNA fragments are simultaneously sequenced.

For maximum flexibility, TruSeq ChIP Library Preparation Kits can be used to prepare samples for single-read or paired-end sequencing, and are compatible with any Illumina sequencing instrument, including MiSeq® and all instruments in the HiSeq® system family.

Table 1: Motif-Finder Analysis of Peaks Identified using TruSeq Library Preparation Kits Compared to ENCODE Reference Peak Data

Name	% Top Peaks with MafK Motif
TruSeq ChIP	95%
ENCODE HELA	92%
ENCODE HES	86%

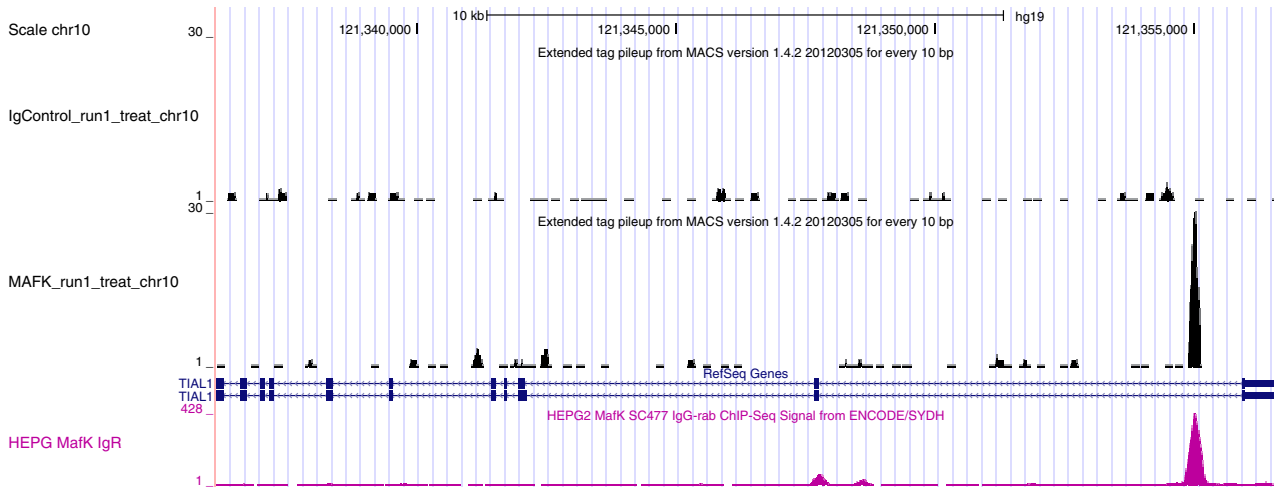
TruSeq Data Quality

Proven TruSeq data quality delivers the most complete and accurate profile of target protein–DNA interactions, enabling an optimal percentage of passing filter reads, percent alignable reads, and coverage uniformity, as well as high sensitivity to detect low-abundance hits.

Robust Multiplex Performance

The TruSeq ChIP Library Preparation Kits provide up to 24 total indexes to increase throughput and consistency without compromising results. The TruSeq universal adapters ligate to sample fragments during library construction, allowing samples to be pooled and individually identified during downstream analysis. This indexing capability improves workflow efficiency and enables robust multiplex sequencing. By enhancing study design flexibility, indexing aids researchers in deriving the most value from each run by efficiently distributing read output based on optimal per-sample read depth requirements.

Figure 3: Peak Finding Output for MafK



TruSeq ChIP Library Preparation Kits enable the generation of libraries across a broad range of study designs. Above is peak data for a negative Ig control, the transcription factor target MafK, and a reference peak for MafK from the ENCODE database.

Flexible Range of Targets

TruSeq ChIP Library Preparation Kits enable libraries to be generated using as little as 5 ng input DNA and provide a high-quality, cost-efficient, and high-throughput solution across a broad array of ChIP study designs. ChIP-Seq is an extremely versatile application that has been successfully applied against a wide range of protein targets, including transcription factors and histones, the building blocks of chromatin. ChIP studies targeting transcription factors are useful in elucidating the specific modulators and signal transduction pathways contributing to disease states, stages of development, or across other conditions, while histone "marks" can be used to better understand how chromatin modifications and local structural changes impact local gene expression activity.

Detecting Peaks Across the Genome

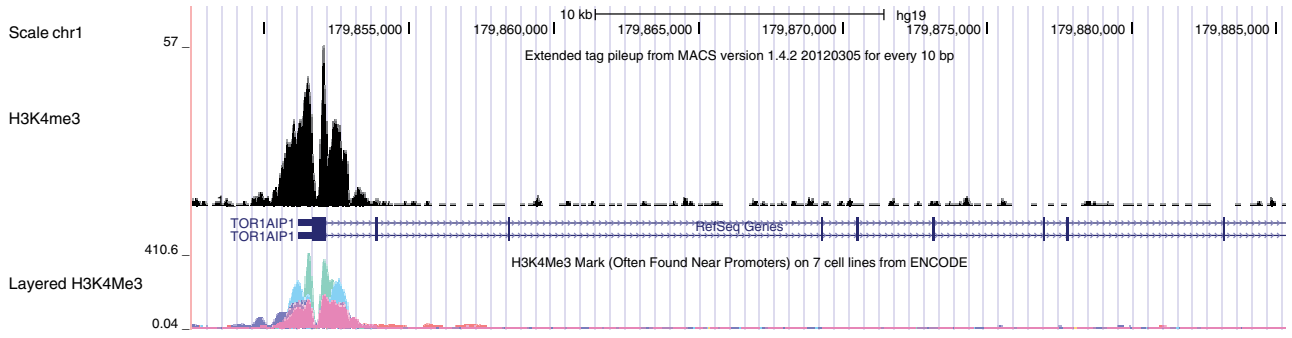
Using the TruSeq ChIP Library Preparation Kit, a library was generated for transcription factor MafK using 5 ng of input DNA (Figure 2) derived from a ChIP performed in HELA cells. Sequencing data were generated using a single MiSeq run. Quality-filtered, BAM output files were then entered into the MACS peak finder software, with the identified peaks then screened for enrichment using MEME motif finder software. Figure 3 illustrates the sensitivity to reliably detect DNA-protein interactions, with a representative, identified peak corresponding to an MafK binding site included in the ENCODE project database. Enrichment for the known, MafK binding motif was detected as expected (Table 1), again in concordance with data generated using MafK peak data available through ENCODE. The ability to robustly detect peaks across the genome with low starting input amounts is critical to ensuring successful ChIP studies.

TruSeq ChIP Library Preparation Kits provide the flexibility to target any protein target of interest, offering a streamlined, cost-efficient solution for studies requiring a broad range of reads per sample including transcription factors (Figure 3), and histone marks, such as H3K4Me3 (Figure 4).

Illumina Sequencing Solutions

TruSeq ChIP Library Preparation Kits are compatible with all Illumina sequencing by synthesis (SBS)–based systems, including the MiSeq and the HiSeq platforms. Offering a revolutionary workflow and unmatched accuracy, MiSeq goes from DNA to data in less than eight hours to support smaller studies. Innovative engineering enables HiSeq systems to process larger numbers of samples quickly and cost-effectively. Data compatibility is ensured whichever system is chosen.

Figure 4: Peak Finding Output for H3K4me3



The peak results for the H3K4me3 target compare favorably with the ENCODE annotation data for this well characterized target, with a representative peak for the histone mark target H3K4me3 and a corresponding ENCODE reference peak.

Summary

TruSeq ChIP Library Preparation Kits offer proven TruSeq accuracy, and a simple, streamlined workflow, enabling highly-multiplexed, cost-effective ChIP sequencing. Supporting analysis of a broad range of targets across the genome even from low sample input, the kits provide a complete, accurate profile of DNA-protein binding interactions and enhanced visibility to the mechanics of gene regulation.

References

1. Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of *in vivo* protein-DNA interactions. Science 316: 1497–1502.
2. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE et al. (2007) High-resolution profiling of histone methylations in the human genome. Cell 129: 823–837.
3. Marban C, Su T, Ferrari R, Li B, Vatakis D, et al. (2011) Genome-wide binding map of the HIV-1 Tat protein to the human genome. PLoS One 6: e26894.
4. Fujiki R, Hashiba W, Sekine H, Yokoyama A, Chikanishi T, et al. (2011) GlcNAcylation of histone H2B facilitates its monoubiquitination. Nature 480: 557–560.
5. Botti E, Spallone G, Moretti F, Marinari B, Pinetti V, et al. (2011) Developmental factor IRF6 exhibits tumor suppressor activity in squamous cell carcinomas. Proc Natl Acad Sci U S A 108: 13710–13715.
6. Bernt KM, Zhu N, Sinha AU, Vempati S, Faber J, et al. (2011) MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. Cancer Cell 20: 66–78.
7. de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, et al. (2011) Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. Nat Struct Mol Biol 18: 977–983.
8. Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, et al. (2011) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 473: 389–393.
9. ENCODE Project Consortium, Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, et al. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 9:e1001046. PMID: 21526222; PMCID: PMC3079585.

Ordering Information

Product	Catalog No.
TruSeq ChIP Library Preparation Kit, Set A (12 indexes, 48 samples)	IP-202-1012
TruSeq ChIP Library Preparation Kit, Set B (12 indexes, 48 samples)	IP-202-1024

Illumina, Inc. • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 770-2012-029 Current as of 12 November 2014

illumina®