Chromatrap

Chromatrap®: a fast, reliable high-throughput ChIP-seq assay for genome-wide protein-DNA analysis

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a technique for genome-wide profiling of DNA-binding proteins and histone modifications. Allied with next-generation sequencing (NGS) technology, Chromatrap® provides a simple and easy to use ChIP format. Compatible with high and low cell numbers, validated on both transcription factor (TF) and epigenetic mark identification in primary and secondary cell lines.

The availability of Chromatrap® ChIP-seq in a 96-well plate allows for high-throughput analysis and, for the first time, provides an unrivalled tool for studying complex, co-ordinated gene regulation and epigenetic mechanisms on a global scale. This application note demonstrates the utility of the Chromatrap® ChIP-seq kit to elucidate the genome-wide binding patterns of nuclear receptor binding events in human adenocarcinoma cells; highlighting the fast multiplex capability alongside massively parallel NGS. **Chromatrap® ChIP-seq assays now enable unbiased, high resolution, low noise amplification for genome-wide understanding of enriched protein-DNA regulatory networks.**

Introduction

Mapping the chromosomal locations of TFs, nucleosomes, histone modifications, chromatin remodelling enzymes, chaperones, and polymerases is a key task in modern biology. The standard methodology for mapping these regulatory protein/DNA interactions is ChIP followed by massively parallel sequencing (ChIP-seq; summarised in Figure 1).

In this application note we present a high quality, reproducible format for ChIP-seq that enables confident analysis of data sets and integrative analysis. When coupled with the ability to conduct multiple target/sample analysis, Chromatrap® ChIP-seq 96 (high-throughput format) enables simultaneous, challenging TF complexes and epigenetic profiles to be deciphered. In this format, Chromatrap® provides a platform for genome-wide, multi-target screening in multiple cell types.

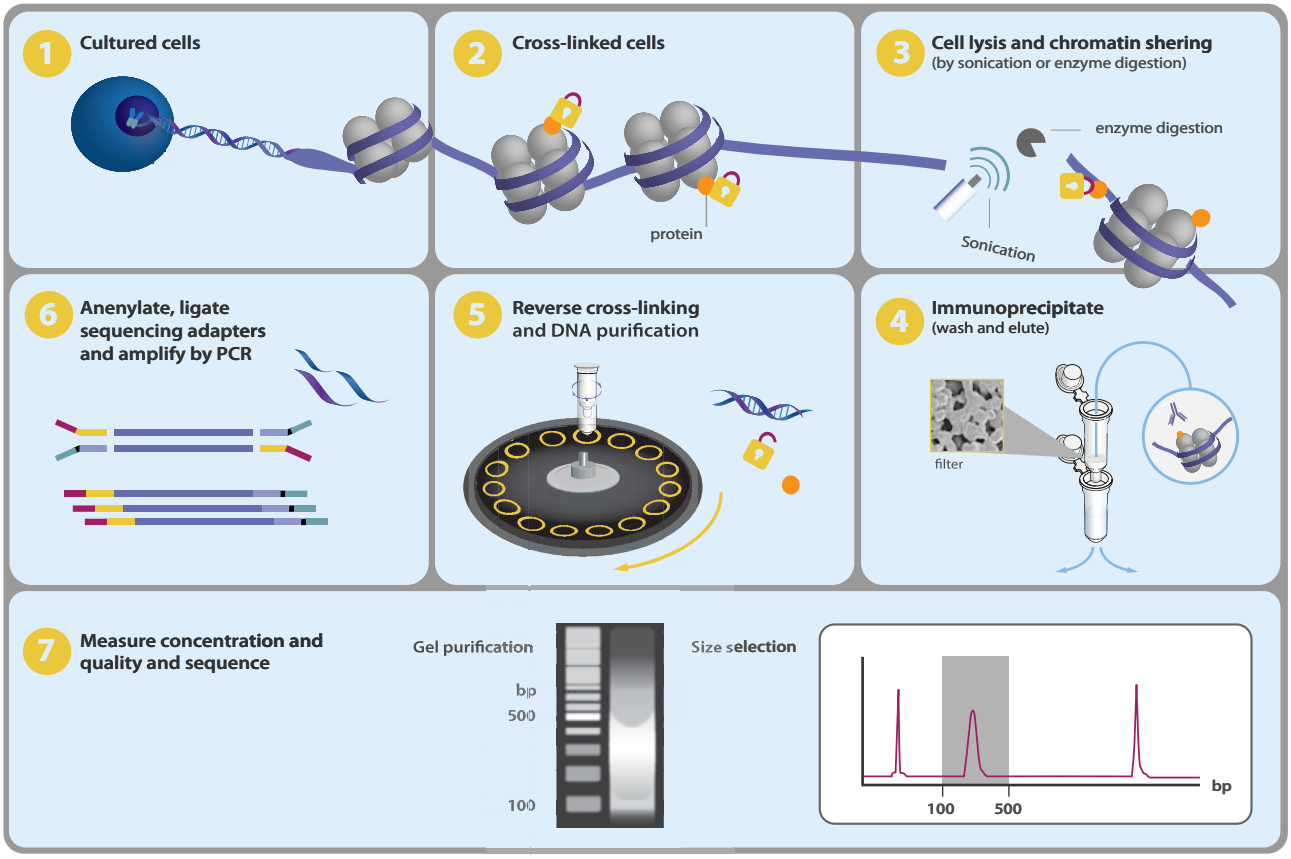


Figure 1 – **Overview of ChIP-seq process**

To demonstrate the ability of Chromatrap® ChIP-seq to selectively and sensitively enrich DNA for NGS, the activity of estrogen receptor (ER) was monitored in a cell line model. **Here we demonstrate both ligand-dependent and -independent binding of ER in cells derived from human endometrial epithelial adenocarcinoma tumours. We also investigated the association of the epigenetic marks H3K4me3 and H3K27me3, these are associated with open and closed chromatin, respectively, in relation to ER binding.**

When coupled with the ability to conduct multiple target/sample analysis, Chromatrap® ChIP-seq in its high-throughput format enables simultaneous, challenging TF complexes and epigenetic profiles to be deciphered. In this format, Chromatrap® now provides a platform for genome-wide, multi-target screening across multiple cell types.

Methodology

1 x 10^7 human endometrial epithelial adenocarcinoma cells were cultured in phenol red-free medium with 10% charcoal-stripped FBS for 48 hours, followed by treatment with 10 nM 17β-estradiol (E2) for 4 hours. Chromatin was prepared according to the **Chromatrap® ChIP-seq spin column protocol** (chromatrap.com/protocols). Prior to sequencing it is important to validate the immunoprecipitation at a specific genomic locus; Chromatrap® provides positive antibody controls and primers for this purpose. In this experiment, validation of the immunoprecipitation enrichment for H3K4me3 and ER antibodies was conducted by qPCR for the GAPDH and XBP1 locus respectively (Figure 2).

Subsequently, immunoprecipitations were performed for sequencing using 30 µg of chromatin and 5 µg of antibody (ER, H3K4me3 or H3K27me3). An equivalent quantity of IgG was also used in a separate IP and used as a background control during data analysis. After immunoprecipitation, chromatin was reverse crosslinked and purified using the illustra GFX PCR DNA and Gel Band Purification Kit and eluted in 50 µl Tris-HCl elution buffer. ChIP DNA was validated using the Agilent Bioanalyzer to verify sample size distribution and concentration (Figure 3).

Samples were quantified using the Qubit 2.0 fluorometer and library preparation was performed according to the Illumina® TruSeq ChIP Sample Prep Kit protocol (IP-102-1012). Illumina® recommends an initial starting quantity of 5-10 ng ChIP DNA. For this study, 25 ng of DNA was used to increase library complexity and minimise the potential for PCR duplications.

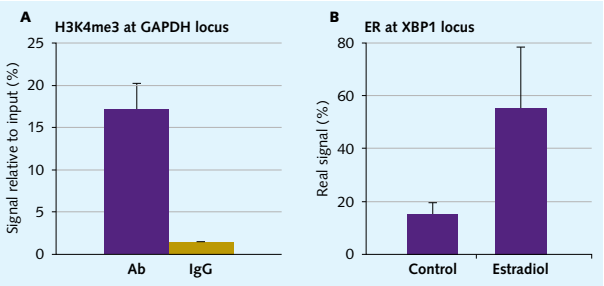


Figure 2 - Chromatin IP validation. *The enrichment of H3K4me3 associated with GAPDH (A) and ER associated with XBP1 (B) was validated prior to library preparation with a small-scale 2 ug IP reaction.*

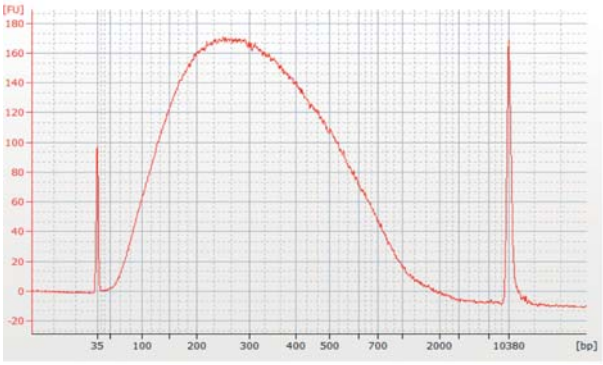


Figure 3. Size selection and library preparation. *The majority of post ChIP sheared chromatin was approximately 200-300 bp.*

Chromatrap® ChIP-seq data analysis – Computational mapping of the sequenced DNA identifies the genomic locations of DNA-binding enzymes, modified histones, chaperones, nucleosomes, and TFs, illuminating the role of these protein-DNA interactions in gene expression and other cellular processes (Mardis et al, 2007). Sequences were mapped to the human reference genome (USCS human genome assembly hg19) and were aligned using Bowtie 2. Data was analysed using MACS or SICER for ER or H3K4me3 and H3K27me3, respectively, using Chromatrap® ChIP-seq Data Analysis software (Figure 5). The corresponding IgG datasets were used as background controls to determine regions of enrichment.

High-throughput sequencing was performed using the Illumina® HiSeq2000 with 1 x 50 bp reads. The quality of sequencing data was analysed by FastQC. On average a Q30 score, which correlates to a 1 in 1000 probability of a base being called incorrectly, applied to >97% of reads for all data sets (Figure 4A). The FastQC scores for the sequencing run and subsequent duplication levels obtained for ERα are shown in figure 4 A, B and C respectively.

In this experiment single end sequencing was used, we have also optimised Chromatrap® with paired end sequencing , which reduces alignment ambiguity to the reference genome, particularly in repetitive regions. For more information on what method of peak calling should be applied to your experiment please refer to our **Top 10 tips for library preparation**.

The Chromatrap® ChIP-seq kit has been optimised for use with Illumina® ChIP-seq library preparation kits and sequencing platforms to deliver high-quality DNA for sequencing. Due to Chromatrap®'s unique solid-phase matrix and the option of a 96-well plate for assays, sample throughput and reproducibility is increased owing to smaller sample volumes, centrifugal wash-steps and a matrix that does not require blocking.

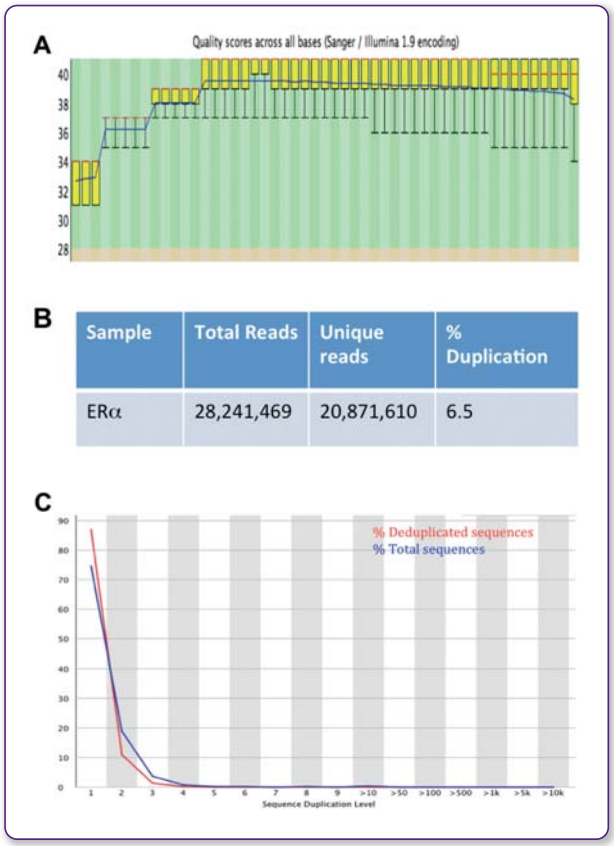


Figure 4 – Sequence quality, total reads and duplication. *FastQC was used to evaluate quality scores across all bases (A), total and unique reads (B), and the sequence duplication level (C).*

Chromatrap® ChIP-seq FastQC sequencing quality parameters – Effective analysis of ChIP-seq data requires sufficient sequencing depth. The required depth depends on the size of the genome and the number and size of the binding sites of the protein. For mammalian TFs which typically have on the order of thousands of binding sites, approximately 20 million reads are required. High abundant TFs (e.g. RNA pol II) and histone marks have many more binding sites and require up to 60 million reads for sufficient coverage in mammalian ChIP-seq [Chen et al 2012].

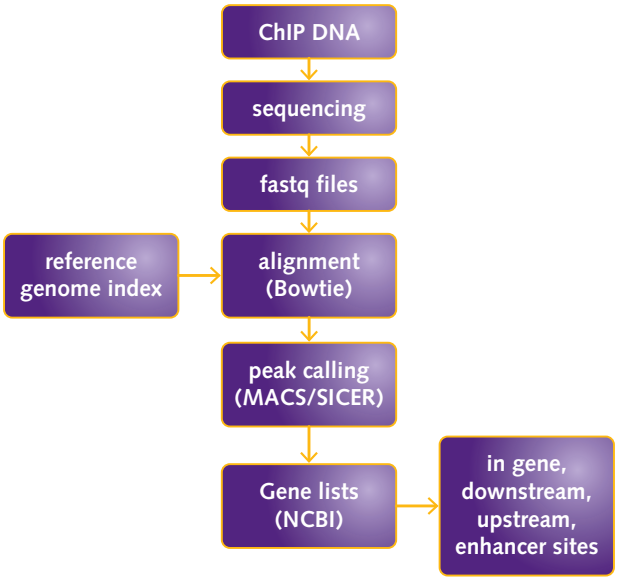


Figure 5 – ChIP-seq data analysis workflow.

Results

Sufficient sequencing depth was obtained, revealing nuclear receptor binding patterns and putative consensus sequences. In the current study, the number of unique reads (74%) obtained from Chromatrap® ChIP-seq allows researchers to ascertain with confidence regions in the genome where their protein of interest is bound and can be used to uncover putative consensus sequences.

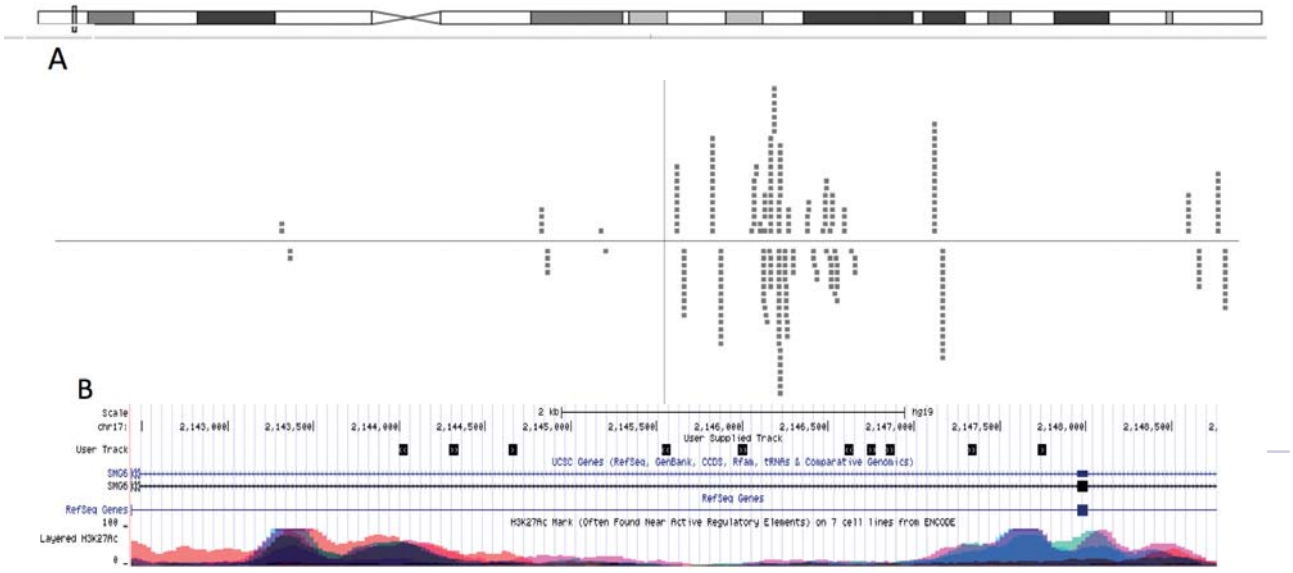
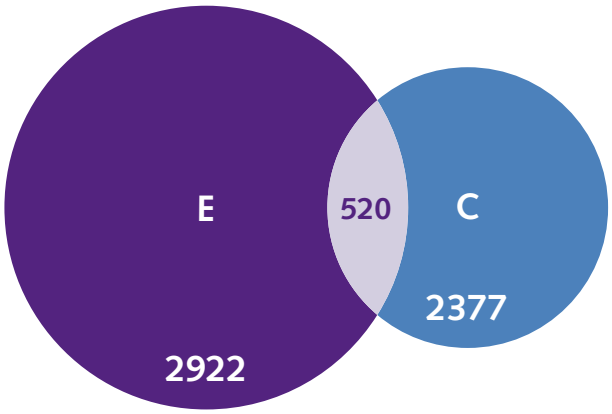


Figure 6 – Graphical visualisation of read alignment. *When aligned to the genome using Bowtie software, MACS-identified peaks can be visualized through genome-browsers. (A) Chipster genome browser view of the chromosomal locus near the SMG6 gene. The dashed lines represents reads, which align to the reference genome that are associated with H3K4me3. (B) UCSC genome browser view of the same chromosomal locus overlayed with H3K27Ac ENCODE data.*

Selectively Enriched Nuclear Receptor Binding

ER's demonstrate a complexity of activation mechanisms and are natural targets for genome wide binding assays. Both ligand dependent and independent binding mechanisms have been shown for ER and this is reflected in the genome wide patterns discovered. In the control, untreated samples, ER was shown to be present at 2377 genes and this enrichment was enhanced following ligand activation (2900, see Figure 7).



	A	B	C	D	E	F	G	H	I	J
1	Chromosom	Associated.G	Gene.Start..I	Gene.End..b	Strand	Description	Peaks	Peak.Pos	Dist.to.Gene.Start	
2		1 RPL23AP21	258019	259027	-1		ISHIKAWACC	downstream	1298, -359, -9636	
3		1 PUSL1	1308582	1311677	1	pseudouridyl	ISHIKAWACC	upstream, uş	-9506, -8258	
4		1 CPSF3L	1311585	1324687	-1		ISHIKAWACC	in gene, in gı	-6503, -5255	
5		1 MIR6727	1312502	1312566	-1		ISHIKAWACC	upstream, uş	-5586, -4338	
6		1 GLTPD1	1324756	1328896	1		ISHIKAWACC	downstream	6668, 7916	
7		1 CAMTA1	6785324	7769706	1		ISHIKAWACC	in gene, in gı	-8296, -52052, -354347,	
8		1 RERE	8352404	8817640	-1	arginine-glut	ISHIKAWACC	in gene, in gı	-238320, -316618	
9		1 KAZN	13893155	15118048	1		ISHIKAWACC	in gene, in gı	-314724, -144349	
10		1 ARHGEF10L	17520197	17697875	1	Rho guanine	ISHIKAWACC	in gene, in gı	-66177, -173465	
11		1 KIAA0319L	35433490	35557670	-1		ISHIKAWACC	in gene, in gı	-88174, -46015	
12		1 PTPRF	43526037	43623672	1	protein tyros	ISHIKAWACC	in gene, in gı	-56122, -52325, -65094	

Figure 7 – Ligand-dependent and -independent ERα binding. *Chromatrap® ChIP-seq resolved uniquely identified genes associated with ER under both estradiol (left), and untreated control conditions (right), as well as genes commonly associated with both (center). In addition, subsequent gene ontology lists were generated that describe unique or commonly associated genes between control and E2 treatment.*

Genes belonging to each treatment group, are compiled into a database that contains information about genomic loci, a description of the gene's function, the start and end position of the gene, and whether the peak occurs upstream, downstream or within the gene itself (Figure 7). Many of the genes associated with the E2 group belong to cell proliferation and metabolic pathways.

Nuclear Receptor function and the epigenetic landscape

H3K4me3 is typically associated with the open chromatin structure of active genes; many of which are constitutively active due to their involvement in essential cellular processes. To demonstrate the utility of high throughput **Chromatrap® ChIP-seq on multiple marks, the simultaneous enrichment of ER and histone methylation marks associated with open and closed chromatin structures was assayed.**

We sought to identify the number of genes that are commonly associated between ER and H3K4me3 or H3K27me3 under basal conditions. Of the 376 genes associated with ER, approximately one third are also commonly found in association with H3K4me3 (Figure 8a). Interestingly, a larger proportion of genes were associated with H3K27me3, an epigenetic modification typically associated with silenced regions (Figure 8b). The number of peaks associated with histone modifications far exceeded those identified for ER.

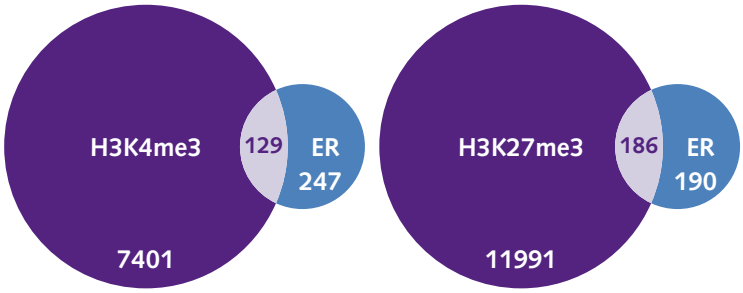


Figure 8. ER recruitment and the histone methylation landscape. *The number of genes associated with H3K4me3 (left) and H3K27me3 (right) when compared with ER binding under basal conditions. Each diagram shows the number of unique and commonly shared genes between each group.*

Summary

New products and technologies are needed to transform genome wide ChIP-seq assays to high-throughput experiments that are capable of mapping multiple marks simultaneously. Chromatrap® ChIP-seq enables further understanding of complex epigenetic and TF mechanisms which control gene regulation. Understanding these mechanisms in both normal physiology and disease is crucial for the design of therapeutic strategies. Furthermore, cohort and metastudies require the need for standardised sample preparation, IP, and analysis techniques, encouraging normalised archiving of large data sets for public use and metadata analysis.

Demonstrated here is the ability of **Chromatrap® ChIP-seq to investigate genome-wide nuclear receptor binding in a complex epigenetic landscape in cell lines expressing multiple isoforms.** Understanding the dynamic nature of TF binding helps to elucidate the complicated transcriptional regulation of genes involved in cancer progression. We were able to identify over 2000 precise genomic locations of direct ER binding and provide information on the association of histone modifications linked with active or silenced chromatin regions.

The Chromatrap® ChIP-seq kit provides the user with high quality DNA to perform multiple library preparations from a single IP, and provides a robust enrichment signal. With chromatin loadings of up to 30 µg, the Chromatrap® ChIP-seq kit provides a dynamic range for sample loading that allows users to load lower amounts of chromatin and minimises antibody requirements. The kit is fully compatible with major sequencing platforms including the Illumina® MiSeq and HiSeq instruments.

References

Mardis, E.R. **ChIP-seq: welcome to the new frontier.** Nature Methods (2007), **4**:613-614.

Chen, Y. **Systematic evaluation of factors influencing ChIP-seq fidelity.** Nature Methods (2012), 9(**6**):609-14.

Chromatrap

European Chromatrap® Technical Support Team Porvair Sciences Ltd
Clywedog Road South Wrexham Industrial Estate Wrexham LL13 9XS UK Tel: +44 (0)1792 606366 Fax: +44 (0)1978 660007

North America Chromatrap® Technical Support Team Porvair Filtration Group Inc.
10190, Maple Leaf Court Ashland Virginia 23005 USA Tel: +1 804 550 1600 Fax: +1 804 550 3262

www.chromatrap.com