Sequencing Methods Review

A review of publications featuring Illumina® Technology

TABLE OF CONTENTS

INTRODUCTION

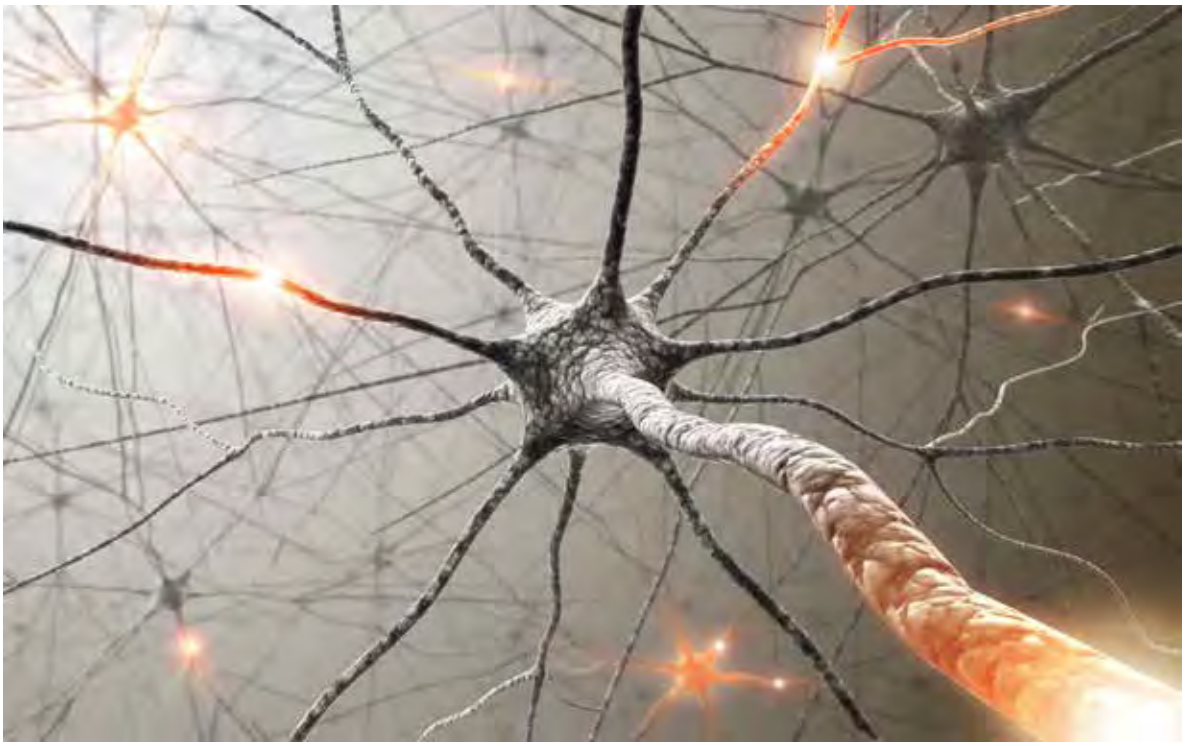
This collection of next-generation sequencing (NGS) sample preparation protocols was compiled from the scientific literature to demonstrate the wide range of scientific questions that can be addressed by Illumina's sequencing by synthesis technology. It is both a tribute to the creativity of the users and the versatility of the technology. We hope it will inspire researchers to use these methods or to develop new ones to address new scientific challenges.

These methods were developed by users, so readers should refer to the original publications for detailed descriptions and protocols.

Have we missed anything? Please contact us if you are aware of a protocol that should be listed.

RNA TRANSCRIPTION

The regulation of RNA transcription and processing directly affects protein synthesis. Proteins, in turn, mediate cellular functions to establish the phenotype of the cell. Dysregulated RNAs are the cause for some diseases and cancers[1,2]. Sequencing RNA provides information about both the abundance and sequence of the RNA molecules. Careful analysis of the results, along with adaptation of the sample preparation protocols, can provide remarkable insight into all the various aspects of RNA processing and control of transcription. Examples of these measures include: post-translational modifications, RNA splicing, RNA bound to RNA binding proteins (RBP), RNA expressed at various stages, unique RNA isoforms, RNA degradation, and regulation of other RNA species[3,4]. Studies of RNA transcription and translation are leading to a better understanding of the implications of RNA production, processing, and regulation for cellular phenotype.



Scientists have discovered a link between long term memory and protein synthesis in brain[5,6].

1 Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

2 Castello A., Fischer B., Hentze M. W. and Preiss T. (2013) RNA-binding proteins in Mendelian disease. Trends Genet 29: 318-327

3 McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

4 Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

5 Davis H. P. and Squire L. R. (1984) Protein synthesis and memory: a review. Psychol Bull 96: 518-559

6 Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

Reviews

Castello A., Fischer B., Hentze M. W. and Preiss T. (2013) RNA-binding proteins in Mendelian disease. Trends Genet 29: 318-327

Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

Law G. L., Korth M. J., Benecke A. G. and Katze M. G. (2013) Systems virology: host-directed approaches]to viral pathogenesis and drug targeting. Nat Rev Microbiol 11: 455-466

Licatalosi D. D. and Darnell R. B. (2010) RNA processing and its regulation: global insights into biological networks. Nat Rev Genet11: 75-87

CHROMATIN ISOLATION BY RNA PURIFICATION (CHIRP-SEQ)

Chromatin isolation by RNA purification (ChIRP-Seq) is a protocol to detect the locations on the genome where non-coding RNAs (ncRNAs), such as long non-coding RNAs (lncRNAs), and their proteins are bound[7]. In this method, samples are first crosslinked and sonicated. Biotinylated tiling oligos are hybridized to the RNAs of interest, and the complexes are captured with streptavidin magnetic beads. After treatment with RNase H the DNA is extracted and sequenced. With deep sequencing the lncRNA/protein interaction site can be determined at single-base resolution.

Pros	Cons
• *Binding sites can be found anywhere on the genome* • *New binding sites can be discovered* • *Specific RNAs of interest can be selected*	• *Nonspecific oligo interactions can lead to misinterpretation of binding sites* • *Chromatin can be disrupted during the preparation stage* • *The sequence of the RNA of interest must be known*

References

Li Z., Chao T. C., Chang K. Y., Lin N., Patil V. S., et al. (2014) The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. Proc Natl Acad Sci U S A 111: 1002-1007

The non-protein–coding parts of the mammalian genome encode thousands of large intergenic non-coding RNAs (lincRNAs). To identify lincRNAs associated with activation of the innate immune response, this study applied custom microarrays and Illumina RNA sequencing for THP1 macrophages. A panel of 159 lincRNAs was found to be differentially expressed following innate activation. Further analysis of the RNA-Seq data revealed that linc1992 was required for expression of many immune-response genes, including cytokines and regulators of TNF-alpha expression.

Illumina Technology: HiSeq 2000®

7 Chu C., Qu K., Zhong F. L., Artandi S. E. and Chang H. Y. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 44: 667-678

Li W., Notani D., Ma Q., Tanasa B., Nunez E., et al. (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498: 516-520

Enhancers are regions of DNA with regulatory function. Through binding of transcription factors and cis-interactions with promoters, target gene expression may be increased. In addition, both lncRNAs and bidirectional ncRNAs may be transcribed on enhancers and are referred to as enhancer RNAs (eRNAs). This study examined eRNA expression in breast cancer cells using a combination of sequencing protocols on HiSeq 2000 (ChIRP-seq, GRO-seq, ChIP-Seq, 3C, 3D-DSL) to discover a global increase in eRNA transcription on enhancers adjacent to E2-upregulated coding genes. These data suggest that eRNAs may play an important role in transcriptional regulation.

Illumina Technology: HiSeq 2000

Chu C., Qu K., Zhong F. L., Artandi S. E. and Chang H. Y. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 44: 667-678

Associated Kits

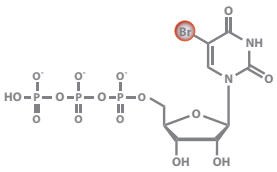
ScriptSeq™ Complete Kit

TruSeq® RNA Sample Prep Kit

TruSeq® Small RNA Sample Prep Kit

GLOBAL RUN-ON SEQUENCING (GRO-SEQ)

Global run-on sequencing (GRO-Seq) maps binding sites of transcriptionally active RNA polymerase II . In this method, active RNA polymerase II[8] is allowed to run on in the presence of Br-UTP. RNAs are hydrolyzed and purified using beads coated with Brd-UTP antibody. The eluted RNA undergoes cap removal and end repair prior to reverse transcription to cDNA. Deep sequencing of the cDNA provides sequences of RNAs that are actively transcribed by RNA polymerase II.



Br-UTP

Pros	Cons
• *Maps position of transcriptionally-engaged RNA polymerases* • *Determines relative activity of transcription sites* • *Detects sense and antisense transcription* • *Detects transcription anywhere on the genome* • *No prior knowledge of transcription sites is needed*	• *The protocol is limited to cell cultures and other artificial systems due to the requirement for incubation in the presence of labeled nucleotides* • *Artifacts may be introduced during the preparation of the nuclei*[9] • *New initiation events may occur during the run-on step* • *Physical impediments may block the polymerases*

References

Heinz S., Romanoski C. E., Benner C., Allison K. A., Kaikkonen M. U., et al. (2013) Effect of natural genetic variation on enhancer selection and function. Nature 503: 487-492

Previous work in epigenetics has proposed a model where lineage-determining transcription factors (LDTF) collaboratively compete with nucleosomes to bind DNA in a cell type–specific manner. In order to determine the sequence variants that guide transcription factor binding, the authors of this paper tested this model in vivo by comparing the SNPs that disrupted transcription factor binding sites in two inbred mouse strains. The authors used GRO-seq in combination with ChIP-seq and RNA-Seq to determine expression and transcription factor binding. The SNPs of the two strains were then classified based on their ability to perturb transcription factor binding and the authors found substantial evidence to support the model.

Illumina Technology: TruSeq RNA Sample Prep Kit, HiSeq 2000

8 Core L. J., Waterfall J. J. and Lis J. T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322: 1845-1848

9 Adelman K. and Lis J. T. (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet 13: 720-731

Jin F., Li Y., Dixon J. R., Selvaraj S., Ye Z., et al. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503: 290-294

Cis-acting regulatory elements in the genome interact with their target gene promoter by transcription factors, bringing the two locations close together in the 3D conformation of the chromatin. In this study the chromosome conformation is examined by a genome-wide analysis method (Hi-C) using the Illumina HiSeq 2000 system. The authors determined over one million long-range chromatin interactions in humanfibroblasts. In addition, they characterized the dynamics of promoter-enhancer contacts after TNF-alpha signaling and discovered pre-existing chromatin looping with the TNF-alpha–responsive enhancers, suggesting the three-dimensional chromatin conformation may be stable over time.

Illumina Technology: HiSeq 2000

Kaikkonen M. U., Spann N. J., Heinz S., Romanoski C. E., Allison K. A., et al. (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol Cell 51: 310-325

Enhancers have been shown to specifically bind lineage-determining transcription factors in a cell-type–specific manner. Toll-like receptor 4 (TLR4) signaling primarily regulates macrophage gene expression through a pre-existing enhancer landscape. In this study the authors used GRO-seq and ChIP-seq to discover that enhancer transcription precedes local mono- and dimethylation of histone H3 lysine 4 (H3K4).

Illumina Technology: Genome Analyzer$_{IIx}$®

Kim Y. J., Greer C. B., Cecchini K. R., Harris L. N., Tuck D. P., et al. (2013) HDAC inhibitors induce transcriptional repression of high copy number genes in breast cancer through elongation blockade. Oncogene 32: 2828-2835

Histone deacetylase inhibitors (HDACI) are a promising class of cancer-repressing drugs. This study investigated the molecular mechanism of HDACI by using GRO-seq in combination with expression analysis. The authors show that HDACI preferentially represses transcription of highly expressed genes which, in cancers, are typically misregulated oncogenes supporting further development of HDACI as a general cancer inhibitor.

Illumina Technology: Genome Analyzer$_{IIx}$, Human Gene Expression—BeadArray; 35 bp reads

Li W., Notani D., Ma Q., Tanasa B., Nunez E., et al. (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498: 516-520

Enhancers are regions of DNA with regulatory function. Through binding of transcription factors and cis-interactions with promoters, target gene expression may be increased. In addition, both lncRNAs and bidirectional ncRNAs may be transcribed on enhancers and are referred to as enhancer RNAs (eRNAs). This study examined eRNA expression in breast cancer cells using a combination of sequencing protocols on HiSeq 2000 (ChIRP-seq, GRO-seq, ChIP-Seq, 3C, 3D-DSL) to discover a global increase in eRNA transcription on enhancers adjacent to E2-upregulated coding genes. These data suggest that eRNAs may play an important role in transcriptional regulation.

Illumina Technology: HiSeq 2000

Saunders A., Core L. J., Sutcliffe C., Lis J. T. and Ashe H. L. (2013) Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. Genes Dev 27: 1146-1158

Drosophila embryogenesis has been intensively studied for the expression patterns of genes corresponding to differentiation of embryonal tissue. In this study, gene regulation was examined using GRO-seq to map the details of RNA polymerase distribution over the genome during early embryogenesis. The authors found that certain groups of genes were more highly paused than others, and that bone morphogenetic protein (BMP) target gene expression requires the pause-inducing negative elongation factor complex (NELF).

Illumina Technology: Genome Analyzer$_{IIx}$

Ji X., Zhou Y., Pandit S., Huang J., Li H., et al. (2013) SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell 153: 855-868

Lam M. T., Cho H., Lesch H. P., Gosselin D., Heinz S., et al. (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498: 511-515

Li P., Spann N. J., Kaikkonen M. U., Lu M., Oh da Y., et al. (2013) NCoR repression of LXRs restricts macrophage biosynthesis of insulin-sensitizing omega 3 fatty acids. Cell 155: 200-214

Chopra V. S., Hendrix D. A., Core L. J., Tsui C., Lis J. T., et al. (2011) The Polycomb Group Mutant esc Leads to Augmented Levels of Paused Pol II in the Drosophila Embryo. Mol Cell 42: 837-844

Hah N., Danko C. G., Core L., Waterfall J. J., Siepel A., et al. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 145: 622-634

Larschan E., Bishop E. P., Kharchenko P. V., Core L. J., Lis J. T., et al. (2011) X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature 471: 115-118

Wang D., Garcia-Bassets I., Benner C., Li W., Su X., et al. (2011) Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474: 390-394

Core L. J., Waterfall J. J. and Lis J. T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322: 1845-1848

Associated Kits

ScriptSeq™ Complete Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA® and Total RNA® Sample Preparation Kit

TruSeq Targeted RNA® Expression Kit

RIBOSOME PROFILING SEQUENCING (RIBO-SEQ)/ARTSEQ™

Active mRNA Translation Sequencing (ARTseq), also called ribosome profiling (Ribo-Seq), isolates RNA that is being processed by the ribosome in order to monitor the translation process[10]. In this method ribosome-bound RNA first undergoes digestion. The RNA is then extracted and the rRNA is depleted. Extracted RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA provides the sequences of RNAs bound by ribosomes during translation. This method has been refined to improve the quality and quantitative nature of the results. Careful attention should be paid to: (1) generation of cell extracts in which ribosomes have been faithfully halted along the mRNA they are translating in vivo; (2) nuclease digestion of RNAs that are not protected by the ribosome followed by recovery of the ribosome-protected mRNA fragments; (3) quantitative conversion of the protected RNA fragments into a DNA library that can be analyzed by deep sequencing[11]. The addition of harringtonine (an alkaloid that inhibits protein biosynthesis) causes ribosomes to accumulate precisely at initiation codons and assists in their detection.

Pros	Cons
• *Reveals a snapshot with the precise location of ribosomes on the RNA*	• *Initiation from multiple sites within a single transcript makes it challenging to define all ORFs*
• *Ribosome profiling more closely reflects the rate of protein synthesis than mRNA levels*	• *Does not provide the kinetics of translational elongation*
• *No prior knowledge of the RNA or ORFs is required*	
• *The whole genome is surveyed*	
• *Can be used to identify protein-coding regions*	

References

Becker A. H., Oh E., Weissman J. S., Kramer G. and Bukau B. (2013) Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. Nat Protoc 8: 2212-2239

A plethora of factors is involved in the maturation of newly synthesized proteins, including chaperones, membrane targeting factors, and enzymes. This paper presents an assay for selective ribosome profiling (SeRP) to determine the interaction of factors with ribosome-nascent chain complexes (RNCs). The protocol is based on Illumina sequencing of ribosome-bound mRNA fragments combined with selection for RNCs associated with the factor of interest.

Illumina Technology: Genome Analyzer$_{IIx}$

10 Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223

11 IIngolia N. T., Lareau L. F. and Weissman J. S. (2011) Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell 147: 789-802

Lee M. T., Bonneau A. R., Takacs C. M., Bazzini A. A., DiVito K. R., et al. (2013) Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature 503: 360-364

In the developmental transition from egg to zygote, the fertilized egg must clear maternal mRNAs and initiate the zygote development program—the zygotic genome activation (ZGA). In this paper, the ZGA was studied in zebrafish using Illumina sequencing to determine the factors that activate the zygotic program. Using a combination of ribosome profiling and mRNA sequencing, the authors identified several hundred genes directly activated by maternal factors, constituting the first wave of zygotic transcription.

Illumina Technology: HiSeq 2000/2500

Stumpf C. R., Moreno M. V., Olshen A. B., Taylor B. S. and Ruggero D. (2013) The translational landscape of the Mammalian cell cycle. Mol Cell 52: 574-582

The regulation of gene expression accounts for the differences seen between different cell types and tissues that share the same genomic information. Regulation may vary over time, and the mechanism and extent is still poorly understood. This study applied Illumina HiSeq technology to sequence total mRNA and total ribosome-occupied mRNA throughout the cell cycle of synchronized HeLa cells to study the translational regulation by ribosome occupancy. The authors identified a large number of mRNAs that undergo significant changes in translation between phases of the cell cycle, and they found 112 mRNAs that were translationally regulated exclusively between specific phases of the cell cycle. The authors suggest translational regulation is a particularly well-suited mechanism for controlling dynamic processes, such as the cell cycle.

Illumina Technology: HiSeq 2000/2500

Wang T., Cui Y., Jin J., Guo J., Wang G., et al. (2013) Translating mRNAs strongly correlate to proteins in a multivariate manner and their translation ratios are phenotype specific. Nucleic Acids Res 41: 4743-4754

It is well known that the abundance of total mRNAs correlates poorly to protein levels. This study set out to analyze the relative abundances of mRNAs, ribosome-nascent chain complex (RNC)-mRNAs, and proteins on a genome-wide scale. A human lung cancer cell line and normal bronchial epithelial cells were analyzed with RNA-seq and the protein abundance measured. The authors created a multivariate linear model showing strong correlation of RNA and protein abundance by integrating the mRNA length as a key factor.

Illumina Technology: Genome Analyzer$_{IIx}$ and HiSeq 2000

Liu B., Han Y. and Qian S. B. (2013) Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. Mol Cell 49: 453-463

Liu X., Jiang H., Gu Z. and Roberts J. W. (2013) High-resolution view of bacteriophage lambda gene expression by ribosome profiling. Proc Natl Acad Sci U S A 110: 11928-11933

Cho J., Chang H., Kwon S. C., Kim B., Kim Y., et al. (2012) LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell 151: 765-777

Fritsch C., Herrmann A., Nothnagel M., Szafranski K., Huse K., et al. (2012) Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res 22: 2208-2218

Gerashchenko M. V., Lobanov A. V. and Gladyshev V. N. (2012) Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. Proc Natl Acad Sci U S A 109: 17394-17399

Han Y., David A., Liu B., Magadan J. G., Bennink J. R., et al. (2012) Monitoring cotranslational protein folding in mammalian cells at codon resolution. Proc Natl Acad Sci U S A 109: 12467-12472

Hsieh A. C., Liu Y., Edlind M. P., Ingolia N. T., Janes M. R., et al. (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature 485: 55-61

Lee S., Liu B., Lee S., Huang S. X., Shen B., et al. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A 109: E2424-2432

Li G. W., Oh E. and Weissman J. S. (2012) The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature 484: 538-541

Stadler M., Artiles K., Pak J. and Fire A. (2012) Contributions of mRNA abundance, ribosome loading, and post- or peri-translational effects to temporal repression of C. elegans heterochronic miRNA targets. Genome Res 22: 2418-2426

Darnell J. C., Van Driesche S. J., Zhang C., Hung K. Y., Mele A., et al. (2011) FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. Cell 146: 247-261

Ingolia N. T., Lareau L. F. and Weissman J. S. (2011) Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell 147: 789-802

Oh E., Becker A. H., Sandikci A., Huber D., Chaba R., et al. (2011) Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell 147: 1295-1308

Han Y., David A., Liu B., Magadan J. G., Bennink J. R., et al. (2012) Monitoring cotranslational protein folding in mammalian cells at codon resolution. Proc Natl Acad Sci U S A 109: 12467-12472

Ingolia N. T. (2010) Genome-wide translational profiling by ribosome footprinting. Methods Enzymol 470: 119-142

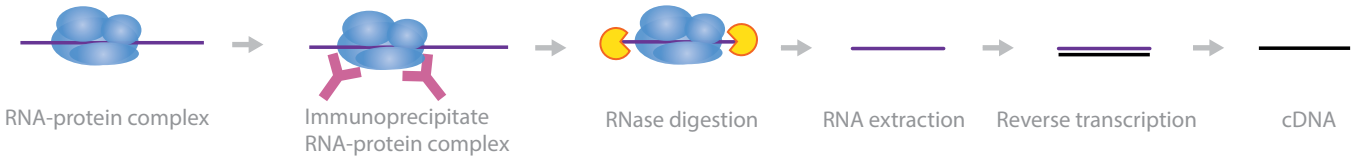
Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero® Kit

RNA IMMUNOPRECIPITATION SEQUENCING (RIP-SEQ)

RNA immunoprecipitation sequencing (RIP-Seq) maps the sites where proteins are bound to the RNA within RNA-protein complexes[12]. In this method, RNA-protein complexes are immunoprecipitated with antibodies targeted to the protein of interest. After RNase digestion, RNA covered by protein is extracted and reverse-transcribed to cDNA. The locations can then be mapped back to the genome. Deep sequencing of cDNA provides single-base resolution of bound RNA.



Pros	Cons
• *Maps specific protein-RNA complexes, such as polycomb-associated RNAs* • *Low background and higher resolution of binding site due to RNase digestion* • *No prior knowledge of the RNA is required* • *Genome-wide RNA screen*	• *Requires antibodies to the targeted proteins* • *Nonspecific antibodies will precipitate nonspecific complexes* • *Lack of crosslinking or stabilization of the complexes may lead to false negatives* • *RNase digestion must be carefully controlled*

References

Kanematsu S., Tanimoto K., Suzuki Y. and Sugano S. (2014) Screening for possible miRNA-mRNA associations in a colon cancer cell line. Gene 533: 520-531

MicroRNAs (miRNAs) are small ncRNAs mediating the regulation of gene expression in various biological contexts, including carcinogenesis. This study examined the putative associations between miRNAs and mRNAs via Argonaute1 (Ago1) or Ago2 immunoprecipitation in a colon cancer cell line. The mRNA sequencing and RIP-seq was performed on an Illumina Genome Analyzer$_{IIx}$ system. From this analysis the authors found specific associations of Ago1 with genes having constitutive cellular functions, whereas putative miRNA-mRNA associations detected with Ago2 IP appeared to be related to signal transduction genes.

Illumina Technology: Genome Analyzer$_{IIx}$

Udan-Johns M., Bengoechea R., Bell S., Shao J., Diamond M. I., et al. (2014) Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. Hum Mol Genet 23: 157-170

Aberrant aggregation of the protein TDP-43 is a key feature of the pathology of amyotrophic lateral sclerosis (ALS). Studying the mechanism of TDP-43 aggregation, this paper presents an analysis of gene expression and RNA-binding partners in human and mouse cell lines. The aggregation of TDP-43 was observed during heat shock and potential interaction partners were identified. The authors suggest TDP-43 shares properties with physiologic prions from yeast, requiring chaperone proteins for aggregation.

Illumina Technology: HiSeq 2000

12 Zhao J., Ohsumi T. K., Kung J. T., Ogawa Y., Grau D. J., et al. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell 40: 939-953

Wang X., Lu Z., Gomez A., Hon G. C., Yue Y., et al. (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505: 117-120

N^6-methyladenosine (m6A) is the most prevalent internal (non-cap) modification present in the messenger RNA of all higher eukaryotes. To understand the role of m6A modification in mammalian cells, the authors of this study applied Illumina sequencing to characterize the YTH domain family 2 (YTHDF2) reader protein regulation of mRNA degradation. The authors performed m6A-seq (MeRIP-Seq), RIP-seq, mRNA-Seq, photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP), and ribosome profiling for HeLa cells on an Illumina HiSeq system with 100 bp single-end reads. They demonstrated that m6A is selectively recognized by YTHDF2, affecting the translation status and lifetime of mRNA.

Illumina Technology: HiSeq 2000; 100 bp single-end reads

Di Ruscio A., Ebralidze A. K., Benoukraf T., Amabile G., Goff L. A., et al. (2013) DNMT1-interacting RNAs block gene-specific DNA methylation. Nature 503: 371-376

DNA methylation is one of the many epigenetic factors that influence the regulation of gene expression. In this paper, the authors show that a novel RNA from the CEBPA gene locus is critical in regulating the local DNA methylation profile, and thus co-influences gene regulation. Using RIP-seq and RNA-Seq on Illumina platforms, the authors showed that this novel RNA binds DNA (cytosine-5)-methyltransferase 1 (DNMT1) and prevents methylation of the CEBPA gene locus.

Illumina Technology: Genome Analyzer$_{IIx}$ and HiSeq 2000

Meyer K. D., Saletore Y., Zumbo P., Elemento O., Mason C. E., et al. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149: 1635-1646

Cernilogar F. M., Onorati M. C., Kothe G. O., Burroughs A. M., Parsi K. M., et al. (2011) Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature 480: 391-395

Salton M., Elkon R., Borodina T., Davydov A., Yaspo M. L., et al. (2011) Matrin 3 binds and stabilizes mRNA. PLoS One 6: e23882

Zhao J., Ohsumi T. K., Kung J. T., Ogawa Y., Grau D. J., et al. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell 40: 939-953

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

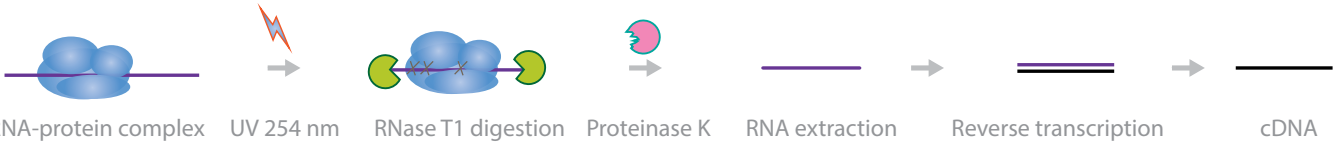
TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

HIGH-THROUGHPUT SEQUENCING OF CLIP CDNA LIBRARY (HITS-CLIP) OR CROSSLINKING AND IMMUNOPRECIPITATION SEQUENCING (CLIP-SEQ)

High-throughput sequencing of CLIP cDNA library (HITS-CLIP) or crosslinking and immunoprecipitation sequencing (CLIP-Seq) maps protein-RNA binding sites *in vivo*[13]. This approach is similar to RIP-Seq, but uses crosslinking to stabilize the protein-RNA complexes. In this method, RNA-protein complexes are UV crosslinked and immunoprecipitated. The protein-RNA complexes are treated with RNase followed by Proteinase K. RNA is extracted and reverse-transcribed to cDNA. Deep sequencing of cDNA provides single-base resolution mapping of protein binding to RNAs.



Pros	Cons
• *Crosslinking stabilizes the protein-target binding* • *UV crosslinking can be carried out in vivo* • *Low background and higher resolution of binding site due to RNase digestion* • *No prior knowledge of the RNA is required* • *Genome-wide RNA screen*	• *Antibodies not specific to the target may precipitate nonspecific complexes* • *UV crosslinking is not very efficient and requires very close protein-RNA interactions* • *Artifacts may be introduced during the crosslinking process*

References

Poulos M. G., Batra R., Li M., Yuan Y., Zhang C., et al. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. Hum Mol Genet 22: 3547-3558

The human muscleblind-like (*MBNL*) genes encode alternative splicing factors essential for development of multiple tissues. In the neuromuscular disease myotonic dystrophy, C(C)UG repeats in RNA inhibit MBNL activity. This paper reports a study of the Mbnl3 protein isoform in a mouse model to determine the function of Mbnl3 in muscle regeneration and muscle function. The authors used an Illumina Genome Analyzer system for RNA-Seq and HITS-CLIP to determine Mbnl3-RNA interaction.

Illumina Technology: Genome Analyzer$_{IIx}$

13 Chi SW, Zang JB, Mele A, Darnell RB; (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460: 479-86

Xu D., Shen W., Guo R., Xue Y., Peng W., et al. (2013) Top3beta is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. Nat Neurosci 16: 1238-1247

Topoisomerases are crucial for solving DNA topological problems, but they have not previously been linked to RNA metabolism. In this study the human topoisomerase 3beta (Top3B), which is known to regulate the translation of mRNAs, was found to bind multiple mRNAs encoded by genes with neuronal functions linked to schizophrenia and autism.

Illumina Technology: Genome Analyzer$_{IIx}$

Charizanis K., Lee K. Y., Batra R., Goodwin M., Zhang C., et al. (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron 75: 437-450

Chi S. W., Hannon G. J. and Darnell R. B. (2012) An alternative mode of microRNA target recognition. Nat Struct Mol Biol 19: 321-327

Riley K. J., Rabinowitz G. S., Yario T. A., Luna J. M., Darnell R. B., et al. (2012) EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. EMBO J 31: 2207-2221

Vourekas A., Zheng Q., Alexiou P., Maragkakis M., Kirino Y., et al. (2012) Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat Struct Mol Biol 19: 773-781

Darnell J. C., Van Driesche S. J., Zhang C., Hung K. Y., Mele A., et al. (2011) FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. Cell 146: 247-261

Polymenidou M., Lagier-Tourenne C., Hutt K. R., Huelga S. C., Moran J., et al. (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci 14: 459-468

Zhang C. and Darnell R. B. (2011) Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol 29: 607-614

McKenna L. B., Schug J., Vourekas A., McKenna J. B., Bramswig N. C., et al. (2010) MicroRNAs control intestinal epithelial differentiation, architecture, and barrier function. Gastroenterology 139: 1654-1664, 1664 e1651

Yano M., Hayakawa-Yano Y., Mele A. and Darnell R. B. (2010) Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. Neuron 66: 848-858

Zhang C., Frias M. A., Mele A., Ruggiu M., Eom T., et al. (2010) Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. Science 329: 439-443

Chi S. W., Zang J. B., Mele A. and Darnell R. B. (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460: 479-486

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

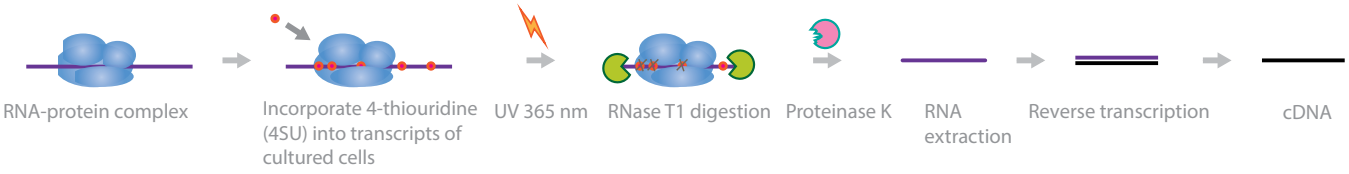
TruSeq Small RNA Sample Prep Kit

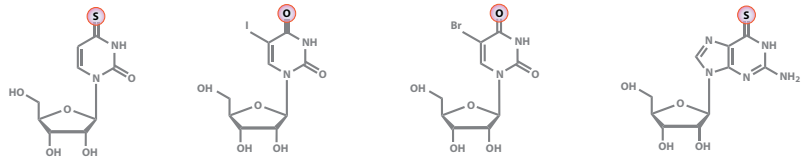
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

PHOTOACTIVATABLE RIBONUCLEOSIDE–ENHANCED CROSSLINKING AND IMMUNOPRECIPITATION (PAR-CLIP)

Photoactivatable ribonucleoside–enhanced crosslinking and immunoprecipitation (PAR-CLIP) maps RNA-binding proteins (RBPs)[14]. This approach is similar to HITS-CLIP and CLIP-Seq, but uses much more efficient crosslinking to stabilize the protein-RNA complexes. The requirement to introduce a photoactivatable ribonucleoside limits this approach to cell culture and *in vitro* systems. In this method, 4-thiouridine (4-SU) and 6-thioguanosine (6-SG) are incorporated into transcripts of cultured cells. UV irradiation crosslinks 4-SU/6-SG–labeled transcripts to interacting RBPs. The targeted complexes are immunoprecipitated and digested with RNase T1, followed by Proteinase K, before RNA extraction. The RNA is reverse-transcribed to cDNA and sequenced. Deep sequencing of cDNA accurately maps RBPs interacting with labeled transcripts.





Photoactivatable ribonucleosides

Pros	Cons
• *Highly accurate mapping of RNA-protein interactions* • *Labeling with 4-SU/6-SG improves crosslinking efficiency*	• *Antibodies not specific to target may precipitate nonspecific complexes* • *Limited to cell culture and in vitro systems*

References

Kaneko S., Bonasio R., Saldana-Meyer R., Yoshida T., Son J., et al. (2014) Interactions between JARID2 and Noncoding RNAs Regulate PRC2 Recruitment to Chromatin. Mol Cell 53: 290-300

JARID2 is an accessory component of Polycomb repressive complex-2 (PRC2) required for the differentiation of embryonic stem cells (ESCs). In this study the molecular role of JARID2 in gene silencing was elucidated using RIP, ChIP, and PAR-CLIP combined with sequencing on an Illumina HiSeq 2000 system. The authors found that Meg3 and other lncRNAs from the Dlk1-Dio3 locus interact with PRC2 via JARID2. These findings suggest a more general mechanism by which lncRNAs contribute to PRC2 recruitment.

Illumina Technology: HiSeq 2000

14 Hafner M., Landgraf P., Ludwig J., Rice A., Ojo T., et al. (2008) Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods 44: 3-12

Liu Y., Hu W., Murakawa Y., Yin J., Wang G., et al. (2013) Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. Sci Rep 3: 2054

In an effort to understand the concert of gene regulation by the circadian rhythm, the authors of this study used a mouse model with a fixed light/dark cycle, to determine genes regulated by variations in body temperature. The authors applied RNA-Seq and PAR-CLIP sequencing on an Illumina Genome Analyzer system to determine Cirbp and Rbm3 as important regulators for the temperature entrained circadian gene expression. They discovered that these two proteins regulate the peripheral clocks by controlling the oscillation of alternative polyadenylation sites.

Illumina Technology: Genome Analyzer®; 76 bp single-end reads

Stoll G., Pietilainen O. P., Linder B., Suvisaari J., Brosi C., et al. (2013) Deletion of TOP3beta, a component of FMRP-containing mRNPs, contributes to neurodevelopmental disorders. Nat Neurosci 16: 1228-1237

Genetic studies, including studies of mRNA-binding proteins, have brought new light to the connection of mRNA metabolism to disease. In this study the authors found the deletion of the topoisomerase 3ß (TOP3ß) gene was associated with neurodevelopmental disorders in the Northern Finnish population. Combining genotyping with immunoprecipitation of mRNA-bound proteins (PAR-CLIP), the authors found that the recruitment of TOP3ß to cytosolic messenger ribonucleoproteins (mRNPs) was coupled to the co-recruitment of FMRP, the disease gene involved in fragile X syndrome mental disorders.

Illumina Technology: Human Gene Expression—BeadArray, Human610-Quad (Infinium GT®), HumanHap300 (Duo/Duo+) (Infinium GT), HumanCNV370-Duo (Infinium GT)

Whisnant A. W., Bogerd H. P., Flores O., Ho P., Powers J. G., et al. (2013) In-depth analysis of the interaction of HIV-1 with cellular microRNA biogenesis and effector mechanisms. MBio 4: e000193

The question of how HIV-1 interfaces with cellular miRNA biogenesis and effector mechanisms has been highly controversial. In this paper, the authors used the Illumina HiSeq 2000 platform for deep sequencing of small RNAs in two different infected cell lines and two types of primary human cells. They unequivocally demonstrated that HIV-1 does not encode any viral miRNAs.

Illumina Technology: TruSeq RNA Sample Prep Kit, HiSeq 2000

Majoros W. H., Lekprasert P., Mukherjee N., Skalsky R. L., Corcoran D. L., et al. (2013) MicroRNA target site identification by integrating sequence and binding information. Nat Methods 10: 630-633

Mandal P. K., Ewing A. D., Hancks D. C. and Kazazian H. H., Jr. (2013) Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. Hum Mol Genet 22: 3730-3748

Hafner M., Lianoglou S., Tuschl T. and Betel D. (2012) Genome-wide identification of miRNA targets by PAR-CLIP. Methods 58: 94-105

Sievers C., Schlumpf T., Sawarkar R., Comoglio F. and Paro R. (2012) Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. Nucleic Acids Res 40: e160

Skalsky R. L., Corcoran D. L., Gottwein E., Frank C. L., Kang D., et al. (2012) The viral and cellular microRNA targetome in lymphoblastoid cell lines. PLoS Pathog 8: e1002484

Uniacke J., Holterman C. E., Lachance G., Franovic A., Jacob M. D., et al. (2012) An oxygen-regulated switch in the protein synthesis machinery. Nature 486: 126-129

Gottwein E., Corcoran D. L., Mukherjee N., Skalsky R. L., Hafner M., et al. (2011) Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines. Cell Host Microbe 10: 515-526

Jungkamp A. C., Stoeckius M., Mecenas D., Grun D., Mastrobuoni G., et al. (2011) In vivo and transcriptome-wide identification of RNA binding protein target sites. Mol Cell 44: 828-840

Kishore S., Jaskiewicz L., Burger L., Hausser J., Khorshid M., et al. (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods 8: 559-564

Lebedeva S., Jens M., Theil K., Schwanhausser B., Selbach M., et al. (2011) Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. Mol Cell 43: 340-352

Mukherjee N., Corcoran D. L., Nusbaum J. D., Reid D. W., Georgiev S., et al. (2011) Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. Mol Cell 43: 327-339

Hafner M., Landthaler M., Burger L., Khorshid M., Hausser J., et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141: 129-141

Hafner M., Landthaler M., Burger L., Khorshid M., Hausser J., et al. (2010) PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

INDIVIDUAL NUCLEOTIDE RESOLUTION CLIP (ICLIP)

Individual nucleotide resolution CLIP (iCLIP) maps protein-RNA interactions similar to HITS-CLIP and PAR-CLIP[15]. This approach includes additional steps to digest the proteins after crosslinking and to map the crosslink sites with reverse transcriptase. In this method specific crosslinked RNA-protein complexes are immunoprecipitated. The complexes are then treated with proteinase K, as the protein crosslinked at the binding site remains undigested. Upon reverse transcription, cDNA truncates at the binding site and is circularized. These circularized fragments are then linearized and PCR-amplified. Deep sequencing of these amplified fragments provides nucleotide resolution of protein-binding site.

Pros	Cons
• *Nucleotide resolution of protein-binding site* • *Avoids the use of nucleases* • *Amplification allows the detection of rare events*	• *Antibodies not specific to target will precipitate nonspecific complexes* • *Non-linear PCR amplification can lead to biases affecting reproducibility* • *Artifacts may be introduced in the circularization step*

References

Broughton J. P. and Pasquinelli A. E. (2013) Identifying Argonaute binding sites in Caenorhabditis elegans using iCLIP. Methods 63: 119-125

The identification of endogenous targets remains an important challenge in understanding miRNA function. New approaches include iCLIP-sequencing, using Illumina sequencing, for high-throughput detection of miRNA targets. In this study the iCLIP protocol was adapted for use in Caenorhabditis elegans to identify endogenous sites targeted by the worm Argonaute protein primarily responsible for miRNA function.

Illumina Technology: Genome Analyzer$_{IIx}$

Zarnack K., Konig J., Tajnik M., Martincorena I., Eustermann S., et al. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell 152: 453-466

Alu elements are a certain type of repeat scattered all over the human genome. Interestingly, Alu elements may be found within gene regions and contain cryptic splice sites. This study investigated the mechanism by which the Alu splice sites are prevented from disrupting normal gene splicing and expression. By using CLIP with Illumina sequencing, the authors profiled mRNAs bound by protein and showed that heterogeneous nuclear riboprotein (hnRNP) C competes with the splicing factor at many genuine and cryptic splice sites. These results suggest hnRNP C acts as a genome-wide protection against transcription disruption by Alu elements.

Illumina Technology: Genome Analyzer$_{IIx}$

15 Konig J., Zarnack K., Rot G., Curk T., Kayikci M., et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17: 909-915

Zund D., Gruber A. R., Zavolan M. and Muhlemann O. (2013) Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat Struct Mol Biol 20: 936-943

UPF1 is a factor involved in nonsense-mediated mRNA decay (NMD). The target binding sites and timing of the binding to target mRNAs has been investigated. In this report the binding sites of UPF1 were studied using transcriptome-wide mapping by CLIP-seq on an Illumina HiSeq 2000 system. The authors show how UPF1 binds RNA before translation and is displaced by translating ribosomes. This observation suggests that the triggering of NMD occurs after the binding of UPF1, presumably through aberrant translation termination.

Illumina Technology: HiSeq 2000

Rogelj B., Easton L. E., Bogu G. K., Stanton L. W., Rot G., et al. (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci Rep 2: 603

Tollervey J. R., Curk T., Rogelj B., Briese M., Cereda M., et al. (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat Neurosci 14: 452-458

Konig J., Zarnack K., Rot G., Curk T., Kayikci M., et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17: 909-915

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

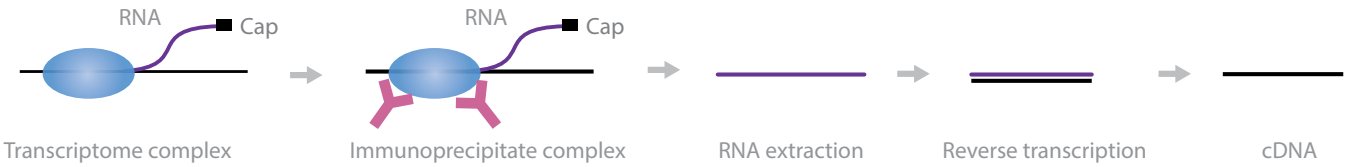
TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

NATIVE ELONGATING TRANSCRIPT SEQUENCING (NET-SEQ)

Native elongating transcript sequencing (NET-Seq) maps transcription through the capture of 3' RNA[16]. In this method the RNA polymerase II elongation complex is immunoprecipitated, and RNA is extracted and reverse-transcribed to cDNA. Deep sequencing of the cDNA allows for 3'-end sequencing of nascent RNA, providing nucleotide resolution at transcription.



Pros	Cons
• *Mapping of nascent RNA-bound protein* • *Transcription is mapped at nucleotide resolution*	• *Antibodies not specific to target will precipitate nonspecific complexes*

References

Larson M. H., Gilbert L. A., Wang X., Lim W. A., Weissman J. S., et al. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc 8: 2180-2196

This paper describes a protocol for selective gene repression based on clustered regularly interspaced palindromic repeats interference (CRISPRi). The protocol provides a simplified approach for rapid gene repression within 1-2 weeks. The method can also be adapted for high-throughput interrogation of genome-wide gene functions and genetic interactions, thus providing a complementary approach to standard RNA interference protocols.

Illumina Technology: HiSeq 2000

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

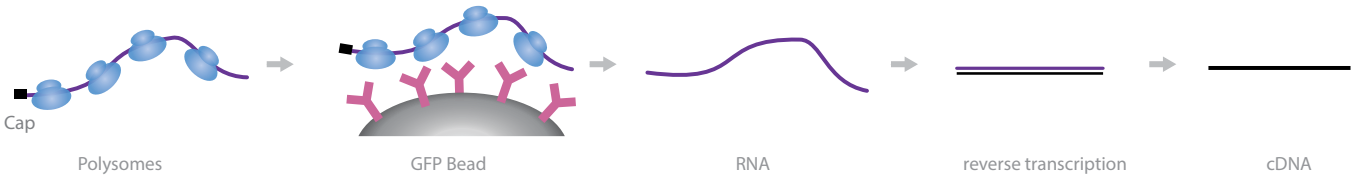
TruSeq Stranded mRNA and Total RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

16 Churchman L. S. and Weissman J. S. (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 469: 368-373

TARGETED PURIFICATION OF POLYSOMAL MRNA (TRAP-SEQ)

Targeted purification of polysomal mRNA (TRAP-Seq) maps translating mRNAs under various conditions[17]. In this method, tagged ribosomal proteins are expressed in cells. The tagged ribosomal proteins are then purified and the RNA isolated. RNAs are reverse-transcribed to cDNA. Deep sequencing of the cDNA provides single-base resolution of translating RNA.



Pros	Cons
• *Allows detection of translating RNAs* • *RNAs translated by specific targeted ribosomes can be assessed* • *No prior knowledge of the RNA is required* • *Genome-wide RNA screen*	• *Not as specific as more recently developed methods, such as Ribo-Seq*

References

Mellen M., Ayata P., Dewell S., Kriaucionis S. and Heintz N. (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151: 1417-1430

Epigenetic markers, such as chromatin-binding factors and modifications to the DNA itself, are important for regulation of gene expression and differentiation. In this study, the DNA methylation 5-hydroxymethylcytosine (5hmC) was profiled in differentiated central nervous system cells *in vivo*. The authors found 5hmC enriched in active genes along with a strong depletion of the alternative methylation 5mC. The authors hypothesize that binding of 5hmC by methyl CpG binding protein 2 (MeCP2) plays a central role in the epigenetic regulation of neural chromatin and gene expression.

Illumina Technology: TruSeq DNA Sample Prep Kit, HiSeq 2000

Associated Kits

ARTseq™ Ribosome Profiling Kit

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

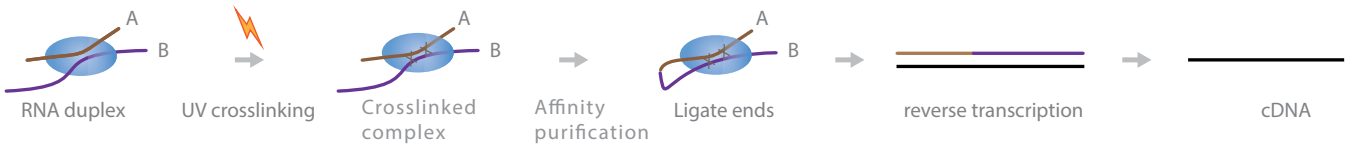
TruSeq Stranded mRNA and Total RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

17 Jiao Y. and Meyerowitz E. M. (2010) Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol 6: 419

CROSSLINKING, LIGATION, AND SEQUENCING OF HYBRIDS (CLASH-SEQ)

Crosslinking, ligation, and sequencing of hybrids (CLASH-Seq) maps RNA-RNA interactions[18]. In this method RNA-protein complexes are UV crosslinked and affinity-purified. RNA-RNA hybrids are then ligated, isolated, and reverse-transcribed to cDNA. Deep sequencing of the cDNA provides high-resolution chimeric reads of RNA-RNA interactions.



Pros	Cons
• *Maps RNA-RNA interactions* • *Performed in vivo*	• *Hybrid ligation may be difficult between short RNA fragments*

References

Kudla G., Granneman S., Hahn D., Beggs J. D. and Tollervey D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci U S A 108: 10010-10015

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

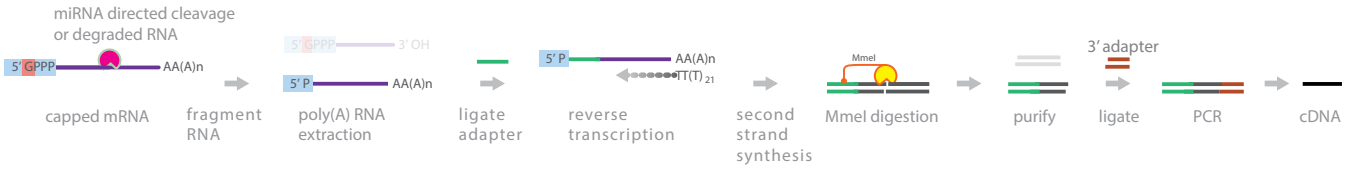
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

18 Kudla G., Granneman S., Hahn D., Beggs J. D. and Tollervey D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci U S A 108: 10010-10015

PARALLEL ANALYSIS OF RNA ENDS SEQUENCING (PARE-SEQ) OR GENOME-WIDE MAPPING OF UNCAPPED TRANSCRIPTS (GMUCT)

Parallel analysis of RNA ends sequencing (PARE-Seq) or genome-wide mapping of uncapped transcripts (GMUCT) maps miRNA cleavage sites. Various RNA degradation processes impart characteristic sequence ends. By analyzing the cleavage sites, the degradation processes can be inferred[19]. In this method, degraded capped mRNA is adapter-ligated and reverse-transcribed. Fragments are then MmeI-digested, purified, 3'-adapter-ligated, and PCR-amplified. Deep sequencing of the cDNA provides information about uncapped transcripts that undergo degradation.



Pros	Cons
• *Maps degrading RNA* • *miRNA cleavage sites are identified* • *No prior knowledge of the target RNA sequence is required*	• *Non-linear PCR amplification can lead to biases, affecting reproducibility* • *Amplification errors caused by polymerases will be represented and sequenced incorrectly*

References

Karlova R, van Haarst JC, Maliepaard C, van de Geest H, Bovy AG, Lammers M, Angenent GC, de Maagd RA; (2013) Identification of microRNA targets in tomato fruit development using high-throughput sequencing and degradome analysis. J Exp Bot 64: 1863-78

The biochemical and genetic processes of fruit development and ripening are of great interest for the food production industry. In this study, the involvement of miRNA in gene regulation was investigated for tomato plants to determine the fruit development processes regulated by miRNA. Using PARE-Seq, the authors identified a total of 119 target genes of miRNAs. Auxin response factors as well as two known ripening regulators were among the identified target genes, indicating an involvement of miRNAs in regulation of fruit ripening.

Illumina Technology: HiSeq 2000

Yang X, Wang L, Yuan D, Lindsey K, Zhang X; (2013) Small RNA and degradome sequencing reveal complex miRNA regulation during cotton somatic embryogenesis. J Exp Bot 64: 1521-36

The authors used PARE-seq to study miRNA expression during cotton somatic embryogenesis. They identified 25 novel miRNAs, as well as their target genes during development.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

19 German M. A., Pillay M., Jeong D. H., Hetawal A., Luo S., et al. (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26: 941-946

Shamimuzzaman M, Vodkin L; (2012) Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. BMC Genomics 13: 310

Bracken CP, Szubert JM, Mercer TR, Dinger ME, Thomson DW, Mattick JS, Michael MZ, Goodall GJ; (2011) Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. Nucleic Acids Res 39: 5658-68

Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood AM, Haugen E, Bracken CP, Rackham O, Stamatoyannopoulos JA, Filipovska A, Mattick JS; (2011) The human mitochondrial transcriptome. Cell 146: 645-58

Associated Kits

TruSeq RNA Sample Prep Kit

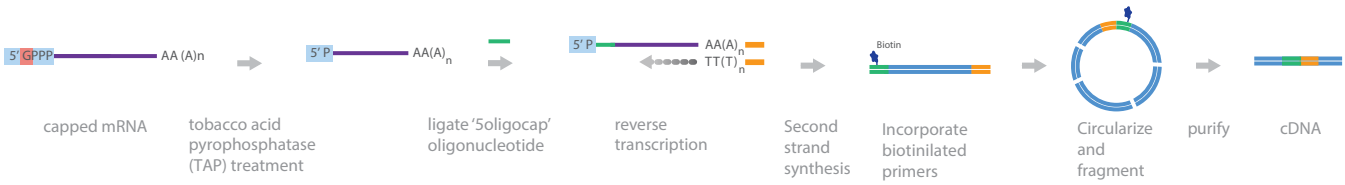
TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

TRANSCRIPT ISOFORM SEQUENCING (TIF-SEQ) OR PAIRED-END ANALYSIS OF TSSS (PEAT)

Transcript isoform sequencing (TIF-Seq)[20] or paired-end analysis of transcription start sites (TSSs) (PEAT)[21] maps RNA isoforms. In this method, the 5' cap is removed with tobacco acid pyrophosphatase (TAP) treatment, then a "5'-oligocap" oligonucleotide is ligated and the RNA is reverse-transcribed. Biotinylated primers are incorporated and the circularized fragment is purified. Deep sequencing of the cDNA provides high-resolution information of the 5' and 3' ends of transcripts.



Pros	Cons
• *Transcript isoforms are identified by 5' and 3' paired-end sequencing*	• *Low-level transcripts may be missed or underrepresented* • *Artifacts may be introduced during the circularization step*

References

Pelechano V., Wei W. and Steinmetz L. M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature 497: 127-131

Identifying gene transcripts by sequencing allows high-throughput profiling of gene expression. However, methods that identify either 5' or 3' transcripts individually do not convey information about the occurrence of transcript isoforms. This paper presents TIF-Seq, a new assay for transcript isoform sequencing. By jointly determining both transcript ends for millions of RNA molecules, this method provides genome-wide detection and annotation of transcript isoforms. The authors demonstrate the TIF-Seq assay for yeast and note that over 26 major transcript isoforms per protein-coding gene were found to be expressed in yeast, suggesting a much higher genome expression repertoire than previously expected.

Illumina Technology: HiSeq 2000

Ni T., Corcoran D. L., Rach E. A., Song S., Spana E. P., et al. (2010) A paired-end sequencing strategy to map the complex landscape of transcription initiation. Nat Methods 7: 521-527

Associated Kits

Ribo-Zero Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

Enzyme Solutions:

Tobacco Acid Pyrophosphatase (TAP)

20 Pelechano V., Wei W. and Steinmetz L. M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature 497: 127-131

21 Ni T., Corcoran D. L., Rach E. A., Song S., Spana E. P., et al. (2010) A paired-end sequencing strategy to map the complex landscape of transcription initiation. Nat Methods 7: 521-527

RNA STRUCTURE

RNA has the ability to form secondary structures that can either promote or inhibit RNA-protein or protein-protein interactions[22,23]. The most diverse secondary and tertiary structures are found in transfer RNAs (tRNAs) and are thought to play a major role in modulating protein translation. RNA structures were first studied in *Tetrahymena thermophilia* using X-ray crystallography, but those studies are inherently cumbersome and limited[24]. Sequencing not only provides information on secondary structures, but it can also determine point mutation effects on RNA structures in a large number of samples. Recent studies have shown that sequencing is a powerful tool to identify RNA structures and determine their significance.

Paramecia species were one of the first model organisms used to study tRNA structure.

Reviews

Lai D., Proctor J. R. and Meyer I. M. (2013) On the importance of cotranscriptional RNA structure formation. RNA 19: 1461-1473

Thapar R., Denmon A. P. and Nikonowicz E. P. (2014) Recognition modes of RNA tetraloops and tetraloop-like motifs by RNA-binding proteins. Wiley Interdiscip Rev RNA 5: 49-67

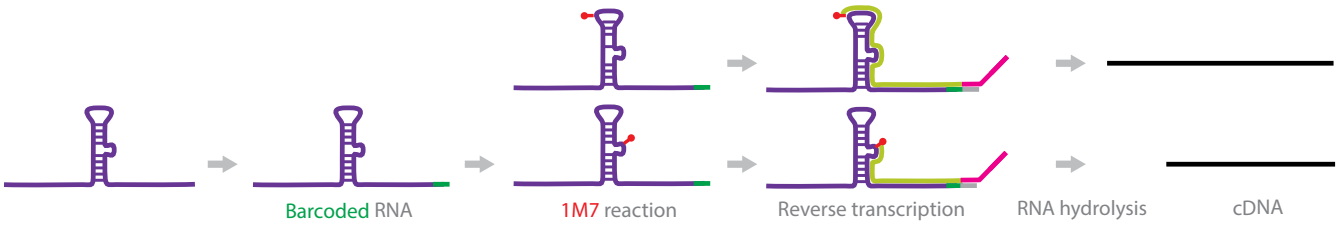
22 Osborne R. J. and Thornton C. A. (2006) RNA-dominant diseases. Hum Mol Genet 15 Spec No 2: R162-169

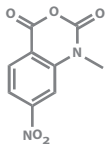
23 Thapar R., Denmon A. P. and Nikonowicz E. P. (2014) Recognition modes of RNA tetraloops and tetraloop-like motifs by RNA-binding proteins. Wiley Interdiscip Rev RNA 5: 49-67

24 Rich A. and RajBhandary U. L. (1976) Transfer RNA: molecular structure, sequence, and properties. Annu Rev Biochem 45: 805-860

SELECTIVE 2'-HYDROXYL ACYLATION ANALYZED BY PRIMER EXTENSION SEQUENCING (SHAPE-SEQ)

Selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)[25] provides structural information about RNA. In this method, a unique barcode is first added to the 3' end of RNA, and the RNA is then allowed to fold under pre-established *in vitro* conditions. The barcoded and folded RNA is treated with a SHAPE reagent, 1M7, that blocks reverse transcription. The RNA is then reverse-transcribed to cDNA. Deep sequencing of the cDNA provides single-nucleotide sequence information for the positions occupied by 1M7. The structural information of the RNA can then be deduced.





1-methyl-7-nitroisatoic anhydride (1M7)

Pros	Cons
• *Provides RNA structural information* • *Multiplexed analysis of barcoded RNAs provides information for multiple RNAs* • *Effect of point mutations on RNA structure can be assessed* • *Alternative to mass spectrometry, NMR, and crystallography*	• *Need positive and negative controls to account for transcriptase drop-off* • *Need pre-established conditions for RNA folding* • *The folding in vitro may not reflect actual folding in vivo*

References

Lucks J. B., Mortimer S. A., Trapnell C., Luo S., Aviran S., et al. (2011) Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Proc Natl Acad Sci U S A 108: 11063-11068

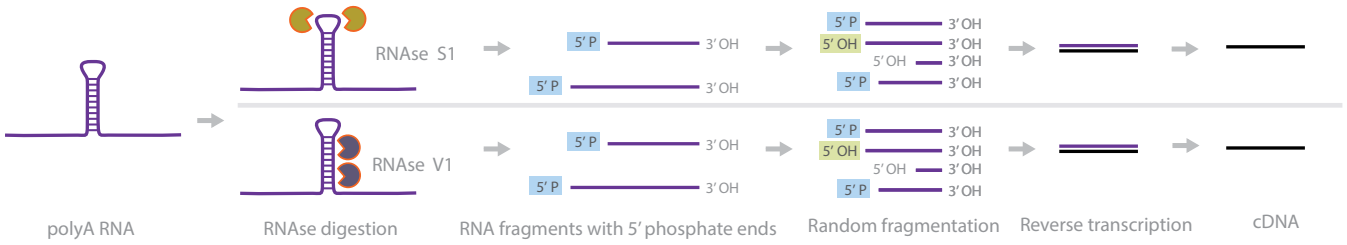
Associated Kits

TruSeq Small RNA Sample Prep Kit

25 Lucks J. B., Mortimer S. A., Trapnell C., Luo S., Aviran S., et al. (2011) Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Proc Natl Acad Sci U S A 108: 11063-11068

PARALLEL ANALYSIS OF RNA STRUCTURE (PARS-SEQ)

Parallel analysis of RNA structure (PARS-Seq)[26] mapping gives information about the secondary and tertiary structure of RNA. In this method RNA is digested with RNases that are specific for double-stranded and single-stranded RNA, respectively. The resulting fragments are reverse-transcribed to cDNA. Deep sequencing of the cDNA provides high-resolution sequences of the RNA. The RNA structure can be deduced by comparing the digestion patterns of the various RNases.



Pros	Cons
• *Provides RNA structural information*	• *Enzyme digestion can be nonspecific*
• *Distinguishes between paired and unpaired bases*	• *Digestion conditions must be carefully controlled*
• *Alternative to mass spectrometry, NMR, and crystallography*	• *RNA can be overdigested*

References

Wan Y, Qu K, Ouyang Z, Chang HY; (2013) Genome-wide mapping of RNA structure using nuclease digestion and high-throughput sequencing. Nat Protoc 8: 849-69

RNA structure is important for RNA function and regulation, and there is growing interest in determining the RNA structure of many transcripts. This is the first paper to describe the PARS protocol. In this method, enzymatic footprinting is coupled with high-throughput sequencing to retrieve information about secondary RNA structure for thousands of RNAs simultaneously.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Associated Kits

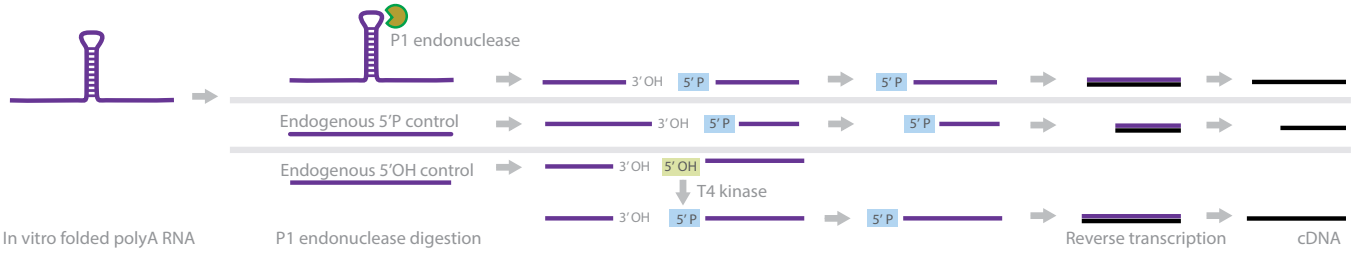
TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

26 Pelechano V., Wei W. and Steinmetz L. M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature 497: 127-131

FRAGMENTATION SEQUENCIN2G (FRAG-SEQ)

Fragmentation sequencing (FRAG-Seq)[27] is a method for probing RNA structure. In this method, RNA is digested using nuclease P1, followed by reverse transcription. Deep sequencing of the cDNA provides high-resolution single-stranded reads, which can be used to determine the structure of RNA by mapping P1 endonuclease digestion sites.



Pros	Cons
• *Simple and fast protocol compared to PARS-seq* • *High throughput* • *Alternative to mass spectrometry, NMR, and crystallography*	• *Need endogenous controls* • *Potential for contamination between samples and controls*

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

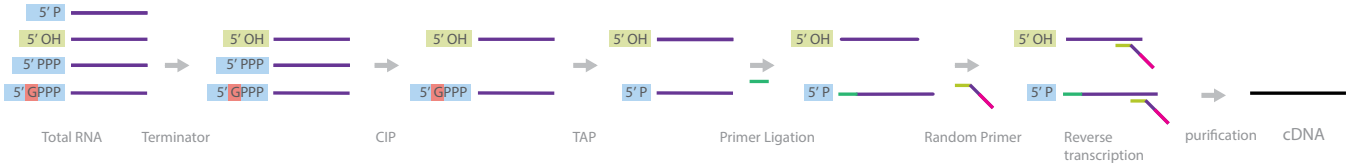
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

27 Underwood J. G., Uzilov A. V., Katzman S., Onodera C. S., Mainzer J. E., et al. (2010) FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat Methods 7: 995-1001

CXXC AFFINITY PURIFICATION SEQUENCING (CAP-SEQ)

CXXC affinity purification sequencing (CAP-Seq)[28] maps the 5' end of RNAs anchored to RNA polymerase II. In this method, RNA transcripts are treated with a terminator, calf intestine alkaline phosphatase (CIP), and then tobacco acid pyrophosphatase (TAP), followed by linker ligation and reverse transcription to cDNA. Deep sequencing of the cDNA provides high-resolution sequences of RNA polymerase II transcripts.



Pros	Cons
• *Maps RNAs anchored to RNA polymerase II*	• *Multiple steps and treatments can lead to loss of material*

References

Farcas A. M., Blackledge N. P., Sudbery I., Long H. K., McGouran J. F., et al. (2012) KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife 1: e00205

DNA methylation occurs naturally throughout the genome, mostly at positions where cytosine is bonded to guanine to form a CpG dinucleotide. Many stretches of CpGs, also called CpG islands, contain a high proportion of unmethylated CpGs. In this study, the unmethylated CpG islands were studied for possible mechanisms favoring the unmethylated sites. Using ChIP-Seq experiments for various transcription factors, the authors showed that CpG islands are occupied by low levels of polycomb repressive complex 1 throughout the genome, potentially making the sites susceptible to polycomb-mediated silencing.

Illumina Technology: HiSeq 2000

28 Illingworth R. S., Gruenewald-Schneider U., Webb S., Kerr A. R., James K. D., et al. (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6: e1001134

Gu W., Lee H. C., Chaves D., Youngman E. M., Pazour G. J., et al. (2012) CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151: 1488-1500

Small RNA molecules account for many different functions in the cell. Piwi-interacting RNAs (piRNAs) represent one type of germline-expressed small RNAs linked to epigenetic programming. This study presents CAP-Seq, an assay developed to characterize the transcription of piRNAs in C. elegans. To their surprise, the authors found that likely piRNA precursors are capped small RNAs that initiate precisely 2 bpupstream of mature piRNAs. In addition, they identified a new class of piRNAs, further adding to the complexity of small RNA molecules.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Clouaire T., Webb S., Skene P., Illingworth R., Kerr A., et al. (2012) Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev 26: 1714-1728

Gendrel A. V., Apedaile A., Coker H., Termanis A., Zvetkova I., et al. (2012) Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive x chromosome. Dev Cell 23: 265-279

Matsushita H., Vesely M. D., Koboldt D. C., Rickert C. G., Uppaluri R., et al. (2012) Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature 482: 400-404

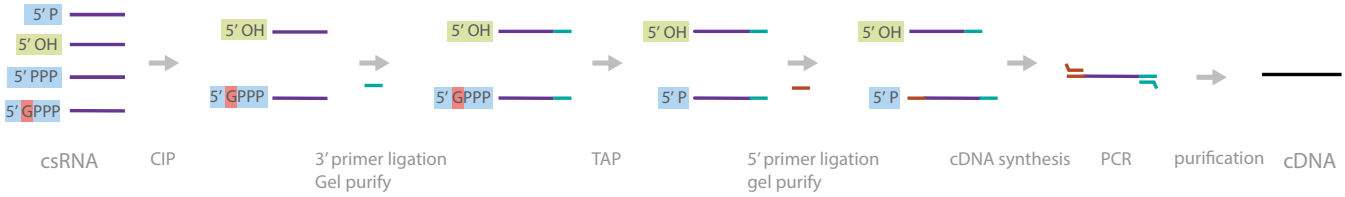
Illingworth R. S., Gruenewald-Schneider U., Webb S., Kerr A. R., James K. D., et al. (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6: e1001134

Associated Kits

- TruSeq RNA Sample Prep Kit
- TruSeq Small RNA Sample Prep Kit
- TruSeq Stranded mRNA and Total RNA® Sample Preparation Kit
- Enzyme Solutions:
  - Tobacco Acid Pyrophosphatase (TAP)
  - Calf Intestinal Phosphatase (CIP)
  - APex Heat-Labile Alkaline Phosphatase

ALKALINE PHOSPHATASE, CALF INTESTINE-TOBACCO ACID PYROPHOSPHATASE SEQUENCING (CIP-TAP)

Alkaline phosphatase, calf intestine-tobacco acid pyrophosphatase sequencing (CIP-TAP) maps capped small RNAs[29]. In this method, RNA is treated with CIP followed by 3'-end linker ligation, then treated with TAP followed by 5'-end linker ligation. The fragments are then reverse-transcribed to cDNA, PCR-amplified, and sequenced. Deep sequencing provides single-nucleotide resolution reads of the capped small RNAs.



Pros	Cons
• *Identifies capped small RNAs missed by CAP-Seq* • *High throughput*	• *Non-linear PCR amplification can lead to biases affecting reproducibility* • *Amplification errors caused by polymerases*

References

Yang L., Lin C., Jin C., Yang J. C., Tanasa B., et al. (2013) lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature 500: 598-602

LncRNAs have recently been indicated to play a role in physiological aspects of cell-type determination and tissue homeostasis. In this paper, the authors applied three sequencing assays (GRO-Seq, ChIRP-Seq, and ChIP-Seq) using the Illumina HiSeq 2000 platform to study expression and epigenetic profiles of prostate cancer cells. The authors found two lncRNAs highly overexpressed and showed that they enhance androgen-receptor-mediated gene activation programs and proliferation of prostate cancer cells.

Illumina Technology: HiSeq 2000

29 Gu W., Lee H. C., Chaves D., Youngman E. M., Pazour G. J., et al. (2012) CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151: 1488-1500

Gu W., Lee H. C., Chaves D., Youngman E. M., Pazour G. J., et al. (2012) CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151: 1488-1500

Small RNA molecules account for many different functions in the cell. Piwi-interacting RNAs (piRNAs) represent one type of germline-expressed small RNAs linked to epigenetic programming. This study presents CAP-Seq, an assay developed to characterize the transcription of piRNAs in C. elegans. To their surprise, the authors found that likely piRNA precursors are capped small RNAs that initiate precisely 2 ntupstream of mature piRNAs. In addition, they identified a new class of piRNAs, further adding to the complexity of small RNA molecules.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

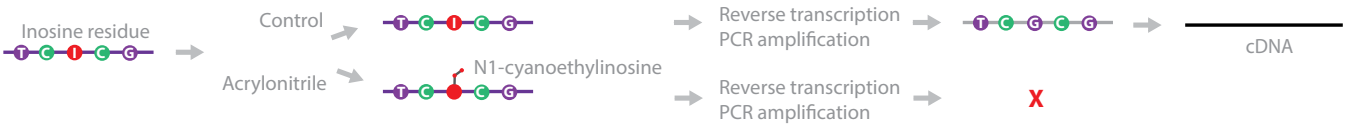
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

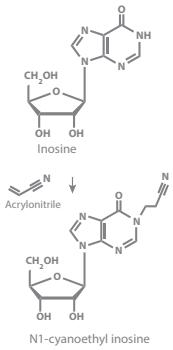
Enzyme Solutions:

- Tobacco Acid Pyrophosphatase (TAP)
- Calf Intestinal Phosphatase (CIP)
- APex Heat-Labile Alkaline Phosphatase

INOSINE CHEMICAL ERASING SEQUENCING (ICE)

Inosine chemical erasing (ICE)[30] identifies adenosine to inosine editing. In this method, RNA is treated with acrylonitrile, while control RNA is untreated. Control and treated RNAs are then reverse-transcribed and PCR-amplified. Inosines in RNA fragments treated with acrylonitrile cannot be reverse-transcribed. Deep sequencing of the cDNA of control and treated RNA provides high-resolution reads of inosines in RNA fragments.





Pros	Cons
• *Mapping of adenosine to inosine editing* • *Can be performed with limited material*	• *Non-linear PCR amplification can lead to biases, affecting reproducibility* • *Amplification errors caused by polymerases will be represented and sequenced incorrectly*

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

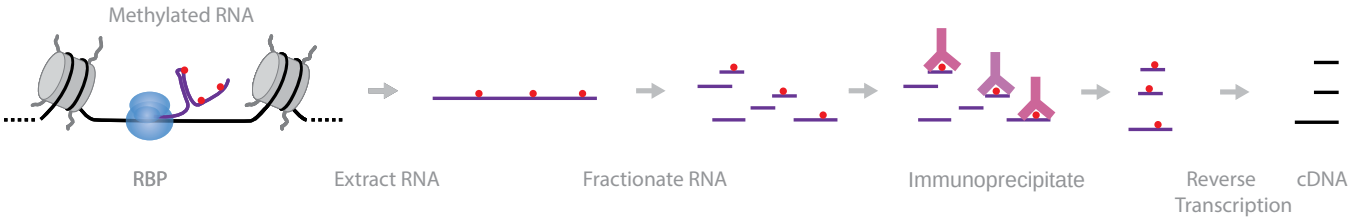
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

30 Sakurai M., Yano T., Kawabata H., Ueda H. and Suzuki T. (2010) Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. Nat Chem Biol 6: 733-740

M^{6}A-SPECIFIC METHYLATED RNA IMMUNOPRECIPITATION SEQUENCING (MERIP-SEQ)

m^{6}A-specific methylated RNA immunoprecipitation with next generation sequencing (MeRIP-Seq)[31] maps m^{6}A methylated RNA. In this method, m^{6}A-specific antibodies are used to immunoprecipitate RNA. RNA is then reverse-transcribed to cDNA and sequenced. Deep sequencing provides high resolution reads of m6A-methylated RNA.



Pros	Cons
• *Maps m^{6}A methylated RNA*	• *Antibodies not specific to target will precipitate nonspecific RNA modifications*

References

Meyer K. D., Saletore Y., Zumbo P., Elemento O., Mason C. E., et al. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149: 1635-1646

In addition to DNA, RNA may also carry epigenetic modifications. Methylation of the N6 position of adenosine (m6A) has been implicated in the regulation of physiological processes. In this study, the authors apply MeRIP-Seq to determine mammalian genes containing m6A in their mRNA. The sites of m6A residues are enriched near stop codons and in 3'-untranslated regions (3'-UTRs), pointing to a non-random distribution and possibly functional relevance of methylated RNA transcripts.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Associated Kits

EpiGnome™ Methyl-Seq® Kit

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

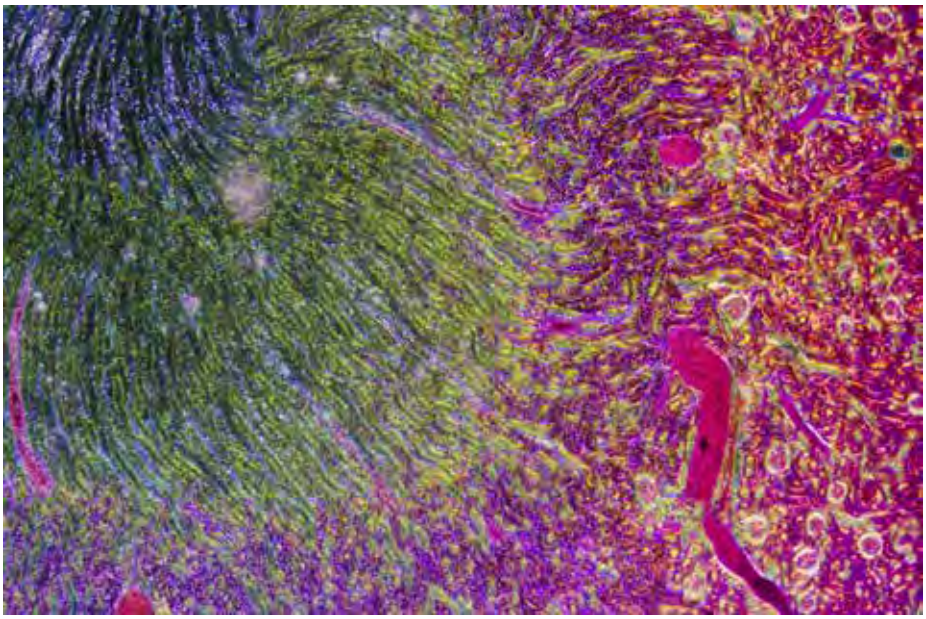
TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

31 Meyer K. D., Saletore Y., Zumbo P., Elemento O., Mason C. E., et al. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149: 1635-1646

LOW-LEVEL RNA DETECTION

Low-level RNA detection refers to both detection of rare RNA molecules in a cell-free environment, such as circulating tumor RNA, or the expression patterns of single cells. Tissues consist of a multitude of different cell types, each with a distinctly different set of functions. Even within a single cell type, the transcriptomes are highly dynamic and reflect temporal, spatial, and cell cycle–dependent changes. Cell harvesting, handling, and technical issues with sensitivity and bias during amplification add an additional level of complexity. To resolve this multi-tiered complexity would require the analysis of many thousands of cells. The use of unique barcodes has greatly increased the number of samples that can be multiplexed and pooled, with little to no decrease in reads associated with each sample. Recent improvements in cell capture and sample preparation will provide more information, faster, and at lower cost[32]. This promises to fundamentally expand our understanding of cell function with significant implications for research and human health[33].



Organs, such as the kidney depicted in this cross-section, consist of a myriad of phenotypically distinct cells. Single-cell transcriptomics can characterize the function of each of these cell types.

Reviews

Blainey P. C. (2013) The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 37: 407-427

Lovett M. (2013) The applications of single-cell genomics. Hum Mol Genet 22: R22-26

Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

Spaethling J. M. and Eberwine J. H. (2013) Single-cell transcriptomics for drug target discovery. Curr Opin Pharmacol 13: 786-790

32 Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

33 Spaethling J. M. and Eberwine J. H. (2013) Single-cell transcriptomics for drug target discovery. Curr Opin Pharmacol 13: 786-790

References

Shalek A. K., Satija R., Adiconis X., Gertner R. S., Gaublomme J. T., et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498: 236-240

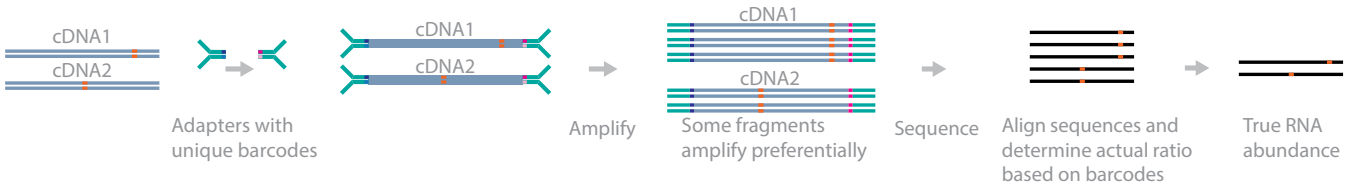
Xue Z., Huang K., Cai C., Cai L., Jiang C. Y., et al. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500: 593-597

Yan L., Yang M., Guo H., Yang L., Wu J., et al. (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 20: 1131-1139

Goetz J. J. and Trimarchi J. M. (2012) Transcriptome sequencing of single cells with Smart-Seq. Nat Biotechnol 30: 763-765

DIGITAL RNA SEQUENCING

Digital RNA sequencing is an approach to RNA-Seq that removes sequence-dependent PCR amplification biases by barcoding the RNA molecules before amplification[34]. RNA is reverse-transcribed to cDNA, then an excess of adapters, each with a unique barcode, is added to the preparation. This barcoded cDNA is then amplified and sequenced. Deep sequencing reads are compared, and barcodes are used to determine the actual ratio of RNA abundance.



Pros	Cons
• *Low amplification bias during PCR* • *Information about abundance of RNA* • *Detection of low-copy–number RNA* • *Single-copy resolution*	• *Some amplification bias still persists* • *Barcodes may miss targets during ligation*

References

Shiroguchi K., Jia T. Z., Sims P. A. and Xie X. S. (2012) Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. Proc Natl Acad Sci U S A 109: 1347-1352

Experimental protocols that include PCR as an amplification step are subject to the sequence-dependent bias of the PCR. For RNA-Seq, this results in difficulties in quantifying expression levels, especially at very low copy numbers. In this study, digital RNA-Seq is introduced as an accurate method for quantitative measurements by appending unique barcode sequences to the pool of RNA fragments. The authors demonstrate how digital RNA-Seq allows transcriptome profiling of Escherichia coli with more accurate and reproducible quantification than conventional RNA-Seq. The efficacy of optimization was estimated by comparison to simulated data.

Illumina Technology: Genome Analyzer$_{IIx}$

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

TruSeq Stranded mRNA and Total RNA Sample Preparation Kit

TruSeq Targeted RNA Expression Kit

34 Shiroguchi K., Jia T. Z., Sims P. A. and Xie X. S. (2012) Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. Proc Natl Acad Sci U S A 109: 1347-1352

WHOLE-TRANSCRIPT AMPLIFICATION FOR SINGLE CELLS (QUARTZ-SEQ)

The Quartz-Seq method optimizes whole-transcript amplification (WTA) of single cells[35]. In this method, a reverse-transcription (RT) primer with a T7 promoter and PCR target is first added to extracted mRNA. Reverse transcription synthesizes first-strand cDNA, after which the RT primer is digested by exonuclease I. A poly(A) tail is then added to the 3' ends of first-strand cDNA, along with a dT primer containing a PCR target. After second-strand generation, a blocking primer is added to ensure PCR enrichment in sufficient quantity for sequencing. Deep sequencing allows for accurate, high-resolution representation of the whole transcriptome of a single cell.

Pros	Cons
• *Single-tube reaction suitable for automation*	• *PCR biases can underrepresent GC-rich templates*
• *Digestion of RT primers by exonuclease I eliminates amplification of byproducts*	• *Amplification errors caused by polymerases will be represented and sequenced incorrectly*
• *Short fragments and byproducts are suppressed during enrichment*	• *Targets smaller than 500 bp are preferentially amplified by polymerases during PCR*

References

Sasagawa Y., Nikaido I., Hayashi T., Danno H., Uno K. D., et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14: R31

Individual cells may exhibit variable gene expression even if they share the same genome. The analysis of single-cell variability in gene expression requires robust protocols with a minimum of bias. This paper presents a novel single-cell RNA-Seq method, Quartz-Seq, based on Illumina sequencing that has a simpler protocol and higher reproducibility and sensitivity than existing methods. The authors implemented improvements in three main areas: 1) they optimized the protocol for suppression of byproduct synthesis; 2) they identified a robust PCR enzyme to allow a single-tube reaction; and 3) they determined optimal conditions for RT and second-strand synthesis.

Illumina Technology: TruSeq RNA Sample Prep Kit, HiSeq 2000

Associated Kits

TruSeq RNA Sample Prep Kit

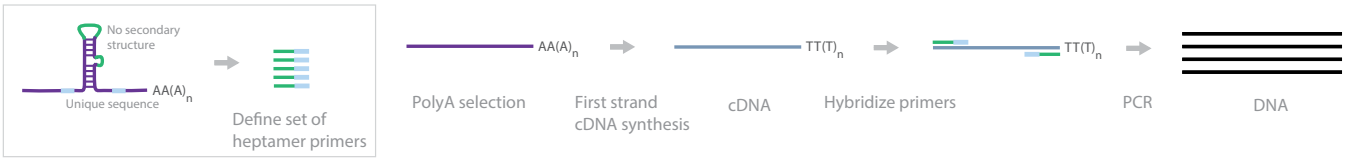
TruSeq Small RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

35 Sasagawa Y., Nikaido I., Hayashi T., Danno H., Uno K. D., et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14: R31

DESIGNED PRIMER–BASED RNA SEQUENCING (DP-SEQ)

Designed Primer–based RNA sequencing (DP-Seq) is a method that amplifies mRNA from limited starting material, as low as 50 pg[36]. In this method, a specific set of heptamer primers are first designed. Enriched poly(A)-selected mRNA undergoes first-strand cDNA synthesis. Designed primers are then hybridized to first-strand cDNA, followed by second strand synthesis and PCR. Deep sequencing of amplified DNA allows for accurate detection of specific mRNA expression at the single-cell level.



Pros	Cons
• *As little as 50 pg of starting material can be used* • *Little transcript-length bias*	• *The sequences of the target areas must be known to design the heptamers* • *Exponential amplification during PCR can lead to primer-dimers and spurious PCR products*[37] • *Some read-length bias*

References

Bhargava V., Ko P., Willems E., Mercola M. and Subramaniam S. (2013) Quantitative transcriptomics using designed primer-based amplification. Sci Rep 3: 1740

Standard amplification of RNA transcripts before sequencing is prone to introduce bias. This paper presents a protocol for selecting a unique subset of primers to target the majority of expressed transcripts in mouse for amplification while preserving their relative abundance. This protocol was developed for Illumina sequencing platforms and the authors show how the protocol yielded high levels of amplification from as little as 50 pg of mRNA, while offering a dynamic range of over five orders of magnitude.

Illumina Technology: Genome Analyzer$_{IIx}$

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

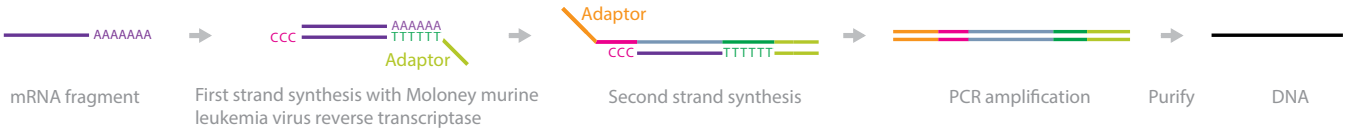
TruSeq Targeted RNA Expression Kit

36 Sasagawa Y., Nikaido I., Hayashi T., Danno H., Uno K. D., et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14: R31

37 Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. (2014) Technical variations in low-input RNA-seq methodologies. Sci Rep 4: 3678

SWITCH MECHANISM AT THE 5' END OF RNA TEMPLATES (SMART-SEQ)

Smart-Seq was developed as a single-cell sequencing protocol with improved read coverage across transcripts[38]. Complete coverage across the genome allows the detection of alternative transcript isoforms and single-nucleotide polymorphisms. In this protocol, cells are lysed and the RNA hybridized to an oligo(dT)-containing primer. The first strand is then created with the addition of a few untemplated C nucleotides. This poly(C) overhang is added exclusively to full-length transcripts. An oligonucleotide primer is then hybridized to the poly(C) overhang and used to synthesize the second strand. Full-length cDNAs are PCR-amplified to obtain nanogram amounts of DNA. The PCR products are purified for sequencing.



Pros	Cons
• *As little as 50 pg of starting material can be used* • *The sequence of the mRNA does not have to be known* • *Improved coverage across transcripts* • *High level of mappable reads*	• *Not strand-specific* • *No early multiplexing*[39] • *Transcript length bias with inefficient transcription of reads over 4 Kb*[40] • *Preferential amplification of high-abundance transcripts* • *The purification step may lead to loss of material* • *Could be subject to strand-invasion bias*[41]

References

Kadkhodaei B., Alvarsson A., Schintu N., Ramsköld D., Volakakis N., et al. (2013) Transcription factor Nurr1 maintains fiber integrity and nuclear-encoded mitochondrial gene expression in dopamine neurons. Proc Natl Acad Sci U S A 110: 2360-2365

Developmental transcription factors important in early neuron differentiation are often found expressed also in the adult brain. This study set out to investigate the development of ventral midbrain dopamine (DA) neurons by studying the transcriptional expression in a mouse model system. By using the Smart-Seq method, which allows sequencing from low amounts of total RNA, the authors could sequence RNA from laser-microdissected DA neurons. Their analysis showed transcriptional activation of the essential transcription factor Nurr1 and its key role in sustaining healthy DA cells.

Illumina Technology: HiSeq 2000, Genomic DNA Sample Prep Kit (FC-102-1001; Illumina)

38 Ramskold D., Luo S., Wang Y. C., Li R., Deng Q., et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30: 777-782

39 Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

40 Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. (2014) Technical variations in low-input RNA-seq methodologies. Sci Rep 4: 3678

41 Tang D. T., Plessy C., Salimullah M., Suzuki A. M., Calligaris R., et al. (2013) Suppression of artifacts and barcode bias in high-throughput transcriptome analyses utilizing template switching. Nucleic Acids Res 41: e44

Marinov G. K., Williams B. A., McCue K., Schroth G. P., Gertz J., et al. (2014) From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. Genome Res 24: 496-510

Recent studies are increasingly discovering cell-to-cell variability in gene expression levels and transcriptional regulation. This study examined the lymphoblastoid cell line GM12878 using the Smart-Seq single-cell RNA-Seq protocol on the Illumina HiSeq 2000 platform to determine variation in transcription among individual cells. The authors determined, through careful quantification, that there aresignificant differences in expression among individual cells, over and above technical variation. In addition, they showed that the transcriptomes from small pools of 30-100 cells approach the information content and reproducibility of contemporary pooled RNA-Seq analysis from large amounts of input material.

Illumina Technology: Nextera DNA® Sample Prep Kit, HiSeq 2000

Shalek A. K., Satija R., Adiconis X., Gertner R. S., Gaublomme J. T., et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498: 236-240

Individual cells can exhibit substantial differences in gene expression, and only recently have genome profiling methods been developed to monitor the expression of single cells. This study applied the Smart-Seq single-cell RNA sequencing on the Illumina HiSeq 2000 platform to investigate heterogeneity in the response of mouse bone marrow–derived dendritic cells (BMDCs) to lipopolysaccharide. The authors found extensive bimodal variation in mRNA abundance and splicing patterns, which was subsequently validated using RNA fluorescence in situ hybridization for select transcripts.

Illumina Technology: HiSeq 2000

Yamaguchi S., Hong K., Liu R., Inoue A., Shen L., et al. (2013) Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. Cell Res 23: 329-339

Mouse primordial germ cells (PGCs) undergo genome-wide DNA methylation reprogramming to reset the epigenome for totipotency. In this study, the dynamics between 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) were characterized using immunostaining techniques and analyzed in combination with transcriptome profiles obtained with Illumina RNA sequencing. The study revealed that the dynamics of 5mC and 5hmC during PGC reprogramming support a model in which DNA demethylation in PGCs occurs through multiple steps, with both active and passive mechanisms. In addition, the transcriptome study suggests that PGC reprogramming may have an important role in the activation of a subset of meiotic and imprinted genes.

Illumina Technology: HiSeq 2000

Ramskold D., Luo S., Wang Y. C., Li R., Deng Q., et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30: 777-782

Yamaguchi S., Hong K., Liu R., Shen L., Inoue A., et al. (2012) Tet1 controls meiosis by regulating meiotic gene expression. Nature 492: 443-447

Associated Kits

Nextera DNA Sample Prep Kit

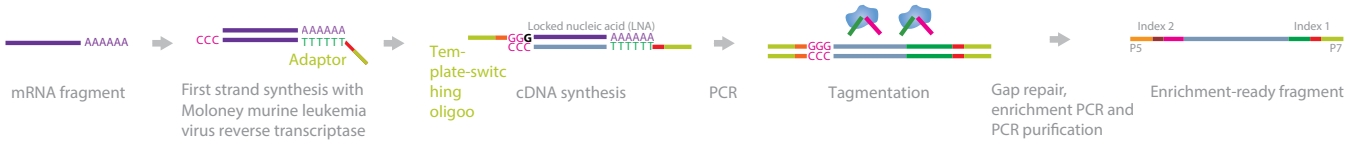
TruSeq RNA Sample Prep Kit

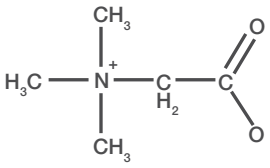
TruSeq Small RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

SWITCH MECHANISM AT THE 5’ END OF RNA TEMPLATES VERSION 2 (SMART-SEQ2)

Smart-Seq2 includes several improvements over the original Smart-Seq protocol[42,43]. The new protocol includes a locked nucleic acid (LNA), an increased $MgCl_2$ concentration, betaine, and elimination of the purification step to significantly improve the yield. In this protocol, single cells are lysed in a buffer that contains free dNTPs and oligo(dT)-tailed oligonucleotides with a universal 5’-anchor sequence. Reverse transcription is performed, which adds 2–5 untemplated nucleotides to the cDNA 3’ end. A template-switching oligo (TSO) is added, carrying two riboguanosines and a modified guanosine to produce a LNA as the last base at the 3’ end. After the first-strand reaction, the cDNA is amplified using a limited number of cycles. Tagmentation is then used to quickly and efficiently construct sequencing libraries from the amplified cDNA.





Betaine

Pros	Cons
• *The sequence of the mRNA does not have to be known* • *As little as 50 pg of starting material can be used* • *Improved coverage across transcripts* • *High level of mappable reads*	• *Not strand-specific* • *No early multiplexing* • *Applicable only to poly(A)+ RNA*

42 Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S., Winberg G., et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10: 1096-1098

43 Picelli S., Faridani O. R., Björklund Å. K., Winberg G., Sagasser S., et al. (2014) Full-length RNA-Seq from single cells using Smart-seq2. Nat. Protocols 9: 171-181

References

Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S., Winberg G., et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10: 1096-1098

Single-cell gene expression analyses hold promise for characterizing cellular heterogeneity, but current methods compromise on the coverage, sensitivity, or throughput. This paper introduces Smart-Seq2 with improved reverse transcription, template switching, and preamplification to increase both yield and length of cDNA libraries generated from individual cells. The authors evaluated the efficacy of the Smart-Seq2 protocol using the Illumina HiSeq 2000 platform and concluded that Smart-Seq2 transcriptome libraries have improved detection, coverage, bias, and accuracy compared to Smart-Seq libraries. In addition, they are generated with off-the-shelf reagents at lower cost.

Illumina Technology: Nextera DNA Sample Prep Kit, HiSeq 2000

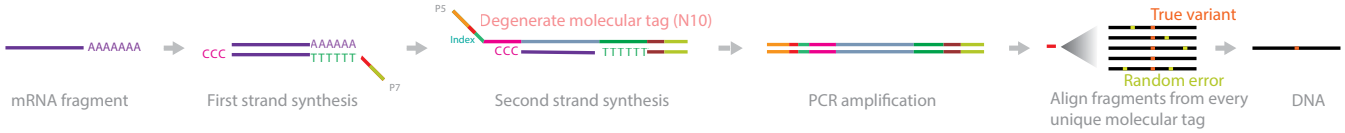
Associated Kits

Nextera DNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

UNIQUE MOLECULAR IDENTIFIERS (UMI)

Unique molecular identifiers (UMI) is a method that uses molecular tags to detect and quantify unique mRNA transcripts[44]. In this method, mRNA libraries are generated by fragmentation and then reverse-transcribed to cDNA. Oligo(dT) primers with specific sequencing linkers are added to cDNA. Another sequencing linker with a 10 bp random label and an index sequence is added to the 5' end of the template, which is amplified and sequenced. Sequencing allows for high-resolution reads, enabling accurate detection of true variants.



Pros	Cons
• *Can sequence unique mRNA transcripts* • *Can be used to detect transcripts occurring at low frequencies* • *Transcripts can be quantified based on sequencing reads specific to each barcode* • *Can be applied to multiple platforms to karyotype chromosomes as well*	• *Targets smaller than 500 bp are preferentially amplified by polymerases during PCR*

References

Islam S., Zeisel A., Joost S., La Manno G., Zajac P., et al. (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods 11: 163-166

Gene expression varies among different tissues, in effect giving rise to different tissue types out of undifferentiated cells; however, expression also varies among different cells in the same tissue. Most assays for measuring gene expression depend on input material from multiple cells, but in this study a method for single-cell RNA sequencing is presented based on Illumina sequencing technology. This technology can be applied to characterize sources of transcriptional noise, or to study expression in early embryos and other sample types where the cell count is naturally limited. One attractive possibility is the application of single-cell sequencing to assess cell type diversity in complex tissues.

Illumina Technology: HiSeq 2000

44 Kivioja T., Vaharautio A., Karlsson K., Bonke M., Enge M., et al. (2012) Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 9: 72-74

Murtaza M., Dawson S. J., Tsui D. W., Gale D., Forshew T., et al. (2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 497: 108-112

Recent studies have shown that genomic alterations in solid cancers can be characterized by sequencing of circulating cell-free tumor DNA released from cancer cells into plasma, representing a non-invasive liquid biopsy. This study describes how this approach was applied using Illumina HiSeq sequencing technology to track the genomic evolution of metastatic cancers in response to therapy. Six patients with breast, ovarian, and lung cancers were followed over 1–2 years. For two cases, synchronous biopsies were also analyzed, confirming genome-wide representation of the tumor genome in plasma and establishing the proof-of-principle of exome-wide analysis of circulating tumor DNA.

Illumina Technology: TruSeq Exome® Enrichment Kit, HiSeq 2000

Kivioja T., Vaharautio A., Karlsson K., Bonke M., Enge M., et al. (2012) Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 9: 72-74

This is the first paper to describe the UMI method and its utility as a tool for sequencing. The authors use UMIs, which make each molecule in a population distinct for genome-scale karyotyping and mRNA sequencing.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Associated Kits

TruSeq RNA Sample Prep Kit

TruSeq Small RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

CELL EXPRESSION BY LINEAR AMPLIFICATION SEQUENCING (CEL-SEQ)

Cell expression by linear amplification sequencing (CEL-Seq) is a method that utilizes barcoding and pooling of RNA to overcome challenges from low input[45]. In this method, each cell undergoes reverse transcription with a unique barcoded primer in its individual tube. After second-strand synthesis, cDNAs from all reaction tubes are pooled, and PCR-amplified. Paired-end deep sequencing of the PCR products allows for accurate detection of sequence derived from sequencing both strands.

Pros	Cons
• *Barcoding and pooling allow for multiplexing and studying many different single cells at a time* • *Cross-contamination is greatly reduced due to using one tube per cell* • *Fewer steps than STRT-Seq* • *Very little read-length bias*[46] • *Strand-specific*	• *Strongly 3' biased*[47] • *Abundant transcripts are preferentially amplified* • *Requires at least 400 pg of total RNA*

References

Hashimshony T., Wagner F., Sher N. and Yanai I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2: 666-673

High-throughput sequencing has allowed for unprecedented detail in gene expression analyses, yet its efficient application to single cells is challenged by the small starting amounts of RNA. This paper presents the CEL-Seq protocol, which uses barcoding, pooling of samples, and linear amplification with one round of in vitro transcription. The assay is designed around a modified version of the Illumina directional RNA protocol and sequencing is done on the Illumina HiSeq 2000 system. The authors demonstrate their method by single-cell expression profiling of early C. *elegans* embryonic development.

Illumina Technology: HiSeq 2000

Associated Kits

TruSeq RNA Sample Prep Kit

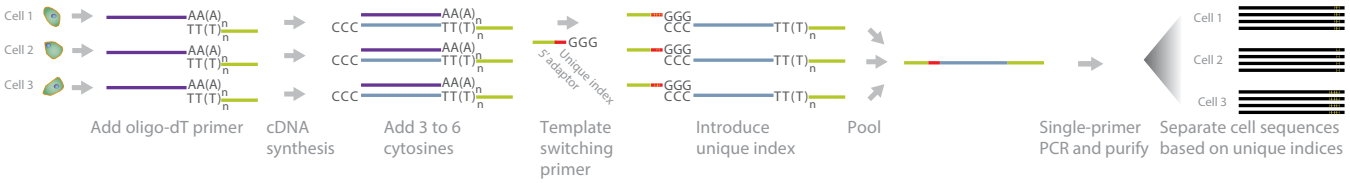
45 Hashimshony T., Wagner F., Sher N. and Yanai I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2: 666-673

46 Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. (2014) Technical variations in low-input RNA-seq methodologies. Sci Rep 4: 3678

47 Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

SINGLE-CELL TAGGED REVERSE TRANSCRIPTION SEQUENCING (STRT-SEQ)

Single-cell tagged reverse transcription sequencing (STRT-Seq) is a method similar to CEL-seq that involves unique barcoding and sample pooling to overcome the challenges of samples with limited material[48]. In this method, single cells are first picked in individual tubes, where first-strand cDNA synthesis occurs using an oligo(dT) primer with the addition of 3–6 cytosines. A helper oligo promotes template switching, which introduces the barcode on the cDNA. Barcoded cDNA is then amplified by single-primer PCR. Deep sequencing allows for accurate transcriptome sequencing of individual cells.



Pros	Cons
• *Barcoding and pooling allows for multiplexing and studying many different single cells at a time* • *Sample handling and the potential for cross-contamination are greatly reduced due to using one tube per cell*	• *PCR biases can underrepresent GC-rich templates* • *Non-linear PCR amplification can lead to biases affecting reproducibility* • *Amplification errors caused by polymerases will be represented and sequenced incorrectly* • *Loss of accuracy due to PCR bias* • *Targets smaller than 500 bp are preferentially amplified by polymerases during PCR*

References

Islam S., Kjallquist U., Moliner A., Zajac P., Fan J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160-1167

Gene expression varies among different tissues, in effect giving rise to different tissue types out of undifferentiated cells; however, expression also varies among different cells in the same tissue. Most assays for measuring gene expression depend on input material from multiple cells, but in this study a method for single-cell RNA sequencing is presented based on Illumina sequencing technology. This technology can be applied to characterize sources of transcriptional noise, or to study expression in early embryos and other sample types where the cell count is naturally limited. One attractive possibility is the application of single-cell sequencing to assess cell type diversity in complex tissues.

Illumina Technology: HiSeq 2000

Associated Kits

TruSeq RNA Sample Prep Kit

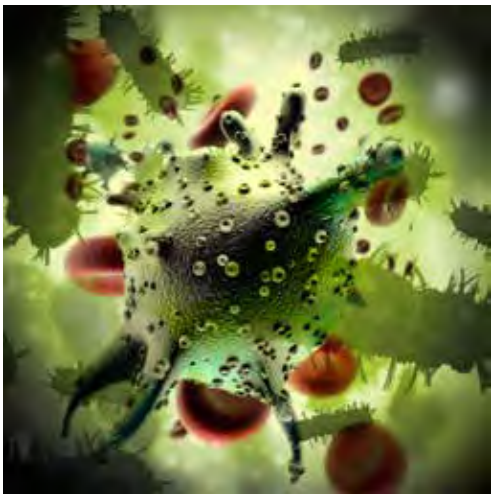
TruSeq Small RNA Sample Prep Kit

TruSeq Targeted RNA Expression Kit

48 Islam S., Kjallquist U., Moliner A., Zajac P., Fan J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160-1167

LOW-LEVEL DNA DETECTION

Single-cell genomics can be used to identify and study circulating tumor cells, cell-free DNA, microbes, uncultured microbes, for preimplantation diagnosis, and to help us better understand tissue-specific cellular differentiation[49, 50]. DNA replication during cell division is not perfect; as a result, progressive generations of cells accumulate unique somatic mutations. Consequently, each cell in our body has a unique genomic signature, which allows the reconstruction of cell lineage trees with very high precision.[51] These cell lineage trees can predict the existence of small populations of stem cells. This information is important for fields as diverse as cancer development[52, 53] preimplantation, and genetic diagnosis. [54, 55]



Single-cell genomics can help characterize and identify circulating tumor cells as well as microbes.

Reviews:

Blainey P. C. (2013) The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 37: 407-427

Lovett M. (2013) The applications of single-cell genomics. Hum Mol Genet 22: R22-26

Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

49 Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

50 Blainey P. C. (2013) The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 37: 407-427

51 Frumkin D., Wasserstrom A., Kaplan S., Feige U. and Shapiro E. (2005) Genomic variability within an organism exposes its cell lineage tree. PLoS Comput Biol 1: e5

52 Navin N., Kendall J., Troge J., Andrews P., Rodgers L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90-94

53 Potter N. E., Ermini L., Papaemmanuil E., Cazzaniga G., Vijayaraghavan G., et al. (2013) Single-cell mutational profiling and clonal phylogeny in cancer. Genome Res 23: 2115-2125

54 Van der Aa N., Esteki M. Z., Vermeesch J. R. and Voet T. (2013) Preimplantation genetic diagnosis guided by single-cell genomics. Genome Med 5: 71

55 Hou Y., Fan W., Yan L., Li R., Lian Y., et al. (2013) Genome analyses of single human oocytes. Cell 155: 1492-1506

Baslan T., Kendall J., Rodgers L., Cox H., Riggs M., et al. (2012) Genome-wide copy number analysis of single cells. Nat Protoc 7: 1024-1041

Böttcher R., Amberg R., Ruzius F. P., Guryev V., Verhaegh W. F., et al. (2012) Using a priori knowledge to align sequencing reads to their exact genomic position. Nucleic Acids Res 40: e125

Kalisky T. and Quake S. R. (2011) Single-cell genomics. Nat Methods 8: 311-314

Navin N. and Hicks J. (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3: 31

Yilmaz S. and Singh A. K. (2011) Single cell genome sequencing. Curr Opin Biotechnol 23: 437-443

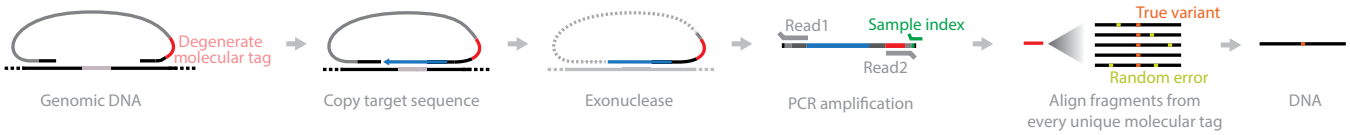
References

Voet T., Kumar P., Van Loo P., Cooke S. L., Marshall J., et al. (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res 41: 6119-6138

Hou Y., Song L., Zhu P., Zhang B., Tao Y., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148: 873-885

SINGLE-MOLECULE MOLECULAR INVERSION PROBES (SMMIP)

The single-molecule molecular inversion probes (smMIP) method uses single-molecule tagging and molecular inversion probes to detect and quantify genetic variations occurring at very low frequencies[56]. In this method, probes are used to detect targets in genomic DNA. After the probed targets are copied, exonuclease digestion leaves the target with a tag, which undergoes PCR amplification and sequencing. Sequencing allows for high-resolution sequence reads of targets, while greater depth allows for better alignment for every unique molecular tag.



Pros	Cons
• *Detection of low-frequency targets* • *Can perform single-cell sequencing or sequencing for samples with very limited starting material*	• *PCR amplification errors* • *PCR biases can underrepresent GC-rich templates* • *Targets smaller than 500 bp are preferentially amplified by polymerases during PCR*

References

Hiatt J. B., Pritchard C. C., Salipante S. J., O'Roak B. J. and Shendure J. (2013) Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. Genome Res 23: 843-854

This is the first paper to describe the smMIP assay, along with its practicality, ability for multiplexing, scaling, and compatibility with desktop sequencing for rapid data collection. The authors demonstrated the assay by resequencing 33 clinically informative cancer genes in 8 cell lines and 45 clinical cancer samples, retrieving accurate data.

Illumina Technology: MiSeq®, HiSeq 2000

Associated Kits

TruSeq Nano DNA® Sample Prep Kit

TruSeq DNA PCR-Free® Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA® Sample Prep Kit

Nextera Rapid Capture Exome/Custom® Enrichment Kit

56 Hiatt J. B., Pritchard C. C., Salipante S. J., O'Roak B. J. and Shendure J. (2013) Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. Genome Res 23: 843-854

MULTIPLE DISPLACEMENT AMPLIFICATION (MDA)

Multiple displacement amplification (MDA) is a method commonly used for sequencing microbial genomes due to its ability to amplify templates larger than 0.5 Mbp, but it can also be used to study genomes of other sizes[57]. In this method, 3'-blocked random hexamer primers are hybridized to the template, followed by synthesis with Phi 29 polymerase. Phi 29 performs strand-displacement DNA synthesis, allowing for efficient and rapid DNA amplification. Deep sequencing of the amplified DNA allows for accurate representation of reads, while sequencing depth provides better alignment and consensus for sequences.

Pros	Cons
• *Templates used for this method can be circular DNA (plasmids, bacterial DNA)* • *Can sequence large templates* • *Can perform single-cell sequencing or sequencing for samples with very limited starting material*	• *Strong amplification bias. Genome coverage as low as ~6%*[58] • *PCR biases can underrepresent GC-rich templates* • *Contaminated reagents can impact results*[59]

References

Embree M., Nagarajan H., Movahedi N., Chitsaz H. and Zengler K. (2013) Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. ISME J

Microbial communities amass a wealth of biochemical processes, and metagenomics approaches are often unable to decipher the key functions of individual microorganisms. This study analyzed a microbial community by first determining the genome sequence of a dominant bacterial member of the genus Smithella, using a single-cell sequencing approach on the Illumina Genome Analyzer. After establishing a working draft genome of Smithella, the authors used low-input metatranscriptomics to determine which genes were active during alkane degradation. The authors then designed a genome-scale metabolic model to integrate the genomic and transcriptomic data.

Illumina Technology: Nextera DNA Sample Prep Kit, MiSeq, Genome Analyzer$_{IIx}$

57 Dean F. B., Nelson J. R., Giesler T. L. and Lasken R. S. (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11: 1095-1099

58 Navin N., Kendall J., Troge J., Andrews P., Rodgers L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90-94

59 Woyke T., Sczyrba A., Lee J., Rinke C., Tighe D., et al. (2011) Decontamination of MDA reagents for single cell whole genome amplification. PLoS ONE 6: e26161

Hou Y., Fan W., Yan L., Li R., Lian Y., et al. (2013) Genome analyses of single human oocytes. Cell 155: 1492-1506

Chromosomal crossover occurs in the oocyte, producing unique combinations of the parent chromosomes in the fertilized egg. This paper presents a protocol for single-cell genome analysis of human oocytes. Using multiple annealing and looping-based amplification cycle (MALBAC)-based sequencing, the authors sequenced triads of the first and second polar bodies from oocyte pronuclei. These pronuclei were derived from the same female egg donors and the authors phased their genomes to determine crossover maps for the oocytes. This breakthrough assay makes important progress toward using whole-genome sequencing for meiosis research and embryo selection for *in vitro* fertilization.

Illumina Technology: HiSeq 2000

McLean J. S., Lombardo M. J., Ziegler M. G., Novotny M., Yee-Greenbaum J., et al. (2013) Genome of the pathogen Porphyromonas gingivalis recovered from a biofilm in a hospital sink using a high-throughput single-cell genomics platform. Genome Res 23: 867-877

Single-cell genomics is becoming an accepted method to capture novel genomes, primarily in marine and soil environments. This study shows, for the first time, that it also enables comparative genomic analysis of strain variation in a pathogen captured from complex biofilm samples in a healthcare facility. The authors present a nearly complete genome representing a novel strain of the periodontal pathogen Porphyromonas gingivalis using the single-cell assembly tool SPAdes.

Illumina Technology: Nextera DNA Sample Prep Kit, Genome Analyzer$_{IIx}$

Seth-Smith H. M., Harris S. R., Skilton R. J., Radebe F. M., Golparian D., et al. (2013) Whole-genome sequences of Chlamydia trachomatis directly from clinical samples without culture. Genome Res 23: 855-866

The use of whole-genome sequencing as a tool to study infectious bacteria is of growing clinical interest. Cultures of Chlamydia trachomatis have, until now, been a prerequisite to obtaining DNA for whole-genome sequencing. Unfortunately, culturing C. trachomatis is a technically demanding and time-consuming procedure. This paper presents IMS-MDA: a new approach combining immunomagnetic separation (IMS) and multiple-displacement amplification (MDA) for whole-genome sequencing of bacterial genomes directly from clinical samples.

Illumina Technology: Genome AnalyzerIIx, HiSeq 2000

Dunowska M., Biggs P. J., Zheng T. and Perrott M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418-424

Wobbly possum disease (WPD) is a fatal neurological disease of the Australian brushtail possum. In this study, the previously unconfirmed mechanism of disease transmission was identified as a novel virus. The identification utilized enrichment for viral DNA followed by sequencing on an Illumina Genome Analyzer.

Illumina Technology: Genome Analyzer$_{IIx}$

Chitsaz H., Yee-Greenbaum J. L., Tesler G., Lombardo M. J., Dupont C. L., et al. (2011) Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. Nat Biotechnol 29: 915-921

Woyke T., Tighe D., Mavromatis K., Clum A., Copeland A., et al. (2010) One bacterial cell, one complete genome. PLoS ONE 5: e10314

Valentim C. L., LoVerde P. T., Anderson T. J. and Criscione C. D. (2009) Efficient genotyping of Schistosoma mansoni miracidia following whole genome amplification. Mol Biochem Parasitol 166: 81-84

Jasmine F., Ahsan H., Andrulis I. L., John E. M., Chang-Claude J., et al. (2008) Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array. Cancer Epidemiol Biomarkers Prev 17: 3499-3508

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

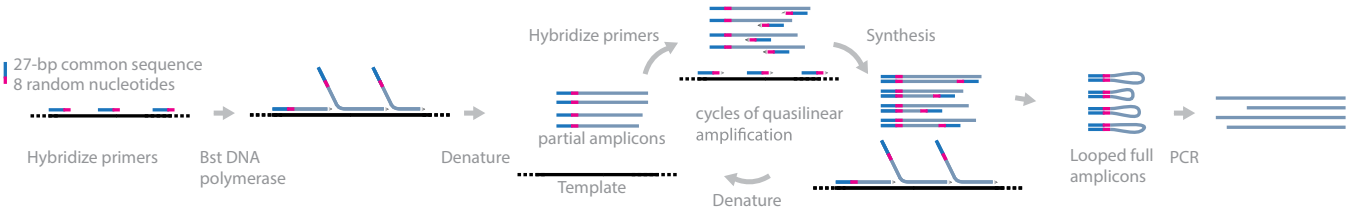
Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

MULTIPLE ANNEALING AND LOOPING–BASED AMPLIFICATION CYCLES (MALBAC)

Multiple annealing and looping–based amplification cycles (MALBAC) is intended to address some of the shortcomings of MDA[60]. In this method, MALBAC primers randomly anneal to a DNA template. A polymerase with displacement activity at elevated temperatures amplifies the template, generating "semi-amplicons." As the amplification and annealing process is repeated, the semi-amplicons are amplified into full amplicons that have a 3' end complimentary to the 5' end. As a result, full-amplicon ends hybridize to form a looped structure, inhibiting further amplification of the looped amplicon, while only the semi-amplicons and genomic DNA undergo amplification. Deep sequencing of the full-amplicon sequences allows for accurate representation of reads, while sequencing depth provides improved alignment for consensus sequences.



Pros	Cons
• *Can sequence large templates* • *Can perform single-cell sequencing or sequencing for samples with very limited starting material* • *Full-amplicon looping inhibits over-representation of templates, reducing PCR bias* • *Can amplify GC-rich regions* • *Uniform genome coverage* • *Lower allele drop-out rate compared to MDA*	• *Polymerase is relatively error prone compared to Phi 29* • *Temperature-sensitive protocol* • *Genome coverage up to ~90%,[61] but some regions of the genome are consistently underrepresented[62]*

References

Hou Y., Fan W., Yan L., Li R., Lian Y., et al. (2013) Genome analyses of single human oocytes. Cell 155: 1492-1506

Chromosomal crossover occurs in the oocyte, producing unique combinations of the parent chromosomes in the fertilized egg. This paper presents a protocol for single-cell genome analysis in human oocytes. Using multiple annealing and looping-based amplification cycle (MALBAC)-based sequencing, the authors sequenced triads of the first and second polar bodies from oocyte pronuclei. These pronuclei were derived from the same female egg donors and the authors phased their genomes to determine crossover maps for the oocytes. This breakthrough assay makes important progress toward using whole-genome sequencing for meiosis research and embryo selection for *in vitro* fertilization.

Illumina Technology: HiSeq 2000

60 Zong C., Lu S., Chapman A. R. and Xie X. S. (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338: 1622-1626

61 Lovett M. (2013) The applications of single-cell genomics. Hum Mol Genet 22: R22-26

62 Lasken R. S. (2013) Single-cell sequencing in its prime. Nat Biotechnol 31: 211-212

Ni X., Zhuo M., Su Z., Duan J., Gao Y., et al. (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110: 21083-21088

There is a great deal of interest in identifying and studying circulating tumor cells (CTCs). Cells from primary tumors enter the bloodstream and can seed metastases. A major barrier to such analysis is low input amounts from single cells, leading to lower coverage. In this study the authors use MALBAC for whole-genome sequencing of single CTCs from patients with lung cancer. They identify copy-number variations that were consistent in patients with the same cancer subtype. Such information about cancers can help identify drug resistance and cancer subtypes, and offers potential for diagnostics, allowing for individualized treatment.

Illumina Technology: MiSeq, HiSeq 2000

Zong C., Lu S., Chapman A. R. and Xie X. S. (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338: 1622-1626

This is the first paper that describes the MALBAC method, which the authors indicate has a higher detection efficiency than the traditional MDA method for single-cell studies. The authors show detection of copy-number variations and single-nucleotide variations of single cancer cells with no false positives.

Illumina Technology: HiSeq 2000

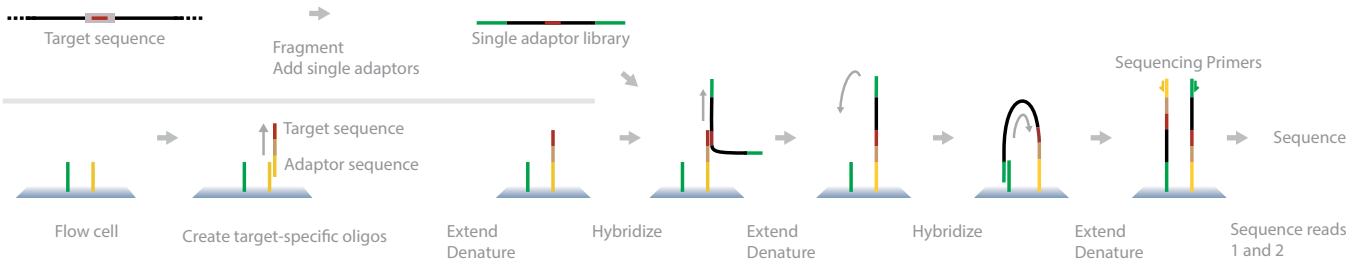
Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

OLIGONUCLEOTIDE-SELECTIVE SEQUENCING (OS-SEQ)

Oligonucleotide-selective sequencing (OS-Seq)[63] was developed to improve targeted resequencing, by capturing and sequencing gene targets directly on the flow cell. In this method target sequences with adapters are used to modify the flow cell primers. Targets in the template are captured onto the flow cell with the modified primers. Further extension, denaturation, and hybridization provide sequence reads for target genes. Deep sequencing provides accurate representation of reads.



Pros	Cons
• *Can resequence multiple targets at a time* • *No gel excision or narrow size purification required* • *Very fast (single-day) protocol* • *Samples can be multiplexed* • *Reduced PCR bias due to removal of amplification steps* • *Avoids loss of material*	• *Primers may interact with similar target sequences, leading to sequence ambiguity*

References

Myllykangas S., Buenrostro J. D., Natsoulis G., Bell J. M. and Ji H. P. (2011) Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing. Nat Biotechnol 29: 1024-1027

As a new method for targeted genome resequencing, the authors present OS-Seq. The method uses a modification of the immobilized lawn of oligonucleotide primers on the flow cell to function as both a capture and sequencing substrate. The method is demonstrated by targeted sequencing of tumor/normal tissue from colorectal cancer.

Illumina Technology: Genome Analyzer$_{IIx}$

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

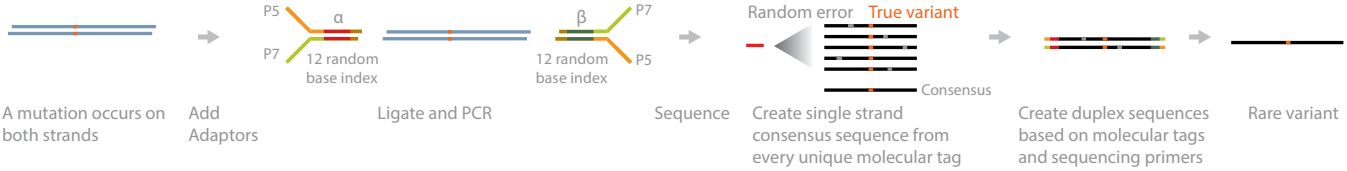
Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

63 Myllykangas S., Buenrostro J. D., Natsoulis G., Bell J. M. and Ji H. P. (2011) Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing. Nat Biotechnol 29: 1024-1027

DUPLEX SEQUENCING (DUPLEX-SEQ)

Duplex sequencing is a tag-based error correction method to improve sequencing accuracy[64]. In this method, adapters (with primer sequences and random 12 bp indices) are ligated onto the template and amplified using PCR. Deep sequencing provides consensus sequence information from every unique molecular tag. Based on molecular tags and sequencing primers, duplex sequences are aligned, determining the true sequence on each DNA strand.



Pros	Cons
• *Very low error rate due to duplex tagging system*	• *PCR amplification errors*
• *PCR amplification errors can be detected and removed from analysis*	• *Non-linear PCR amplification can lead to biases affecting reproducibility*
• *No additional library preparation steps after addition of adapters*	• *PCR biases can underrepresent GC-rich templates*

References

Kennedy S. R., Salk J. J., Schmitt M. W. and Loeb L. A. (2013) Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. PLoS Genet 9: e1003794

Studies of mitochondrial DNA (mtDNA) mutations have been limited due to technical limitations of the protocols applied. In this paper, the authors present a highly sensitive Duplex-Seq method, based on the HiSeq platform, which can detect a single mutation among >107 wild-type molecules. The authors applied the method to study the accumulation of mutations in mtDNA over the course of 80 years of life. Their results show that the mutation spectra of brain tissue of old compared to young individuals are dominated by transition mutations and not G to T mutations, which are the characteristic mutations caused by oxidative damage.

Illumina Technology: HiSeq 2000/2500; 101 bp paired-end reads

Schmitt M. W., Kennedy S. R., Salk J. J., Fox E. J., Hiatt J. B., et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci U S A 109: 14508-14513

The authors propose a tag-based error correction method to improve sequencing accuracy, especially in heterogeneous samples. The method allows double-stranded DNA sequence read collection, proving mutation status on both strands. The method is demonstrated by sequencing M13mp2 DNA. This method is proposed to be useful for assessing mutations due to DNA damage, as well as the determining the mutational status of genes on both DNA strands.

Illumina Technology: HiSeq 2000

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

64 Schmitt M. W., Kennedy S. R., Salk J. J., Fox E. J., Hiatt J. B., et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci U S A 109: 14508-14513

DNA methylation and hydroxymethylation are involved in development, X-chromosome inactivation, cell differentiation, tissue-specific gene expression, plant epigenetic variation, imprinting, cancers, and diseases[65,66,67,68]. Methylation usually occurs at the 5' position of cytosines and plays a crucial role in gene regulation and chromatin remodeling.

The active agouti gene in mice codes for yellow coat color. When pregnant mice with the active agouti gene are fed a diet rich in methyl donors, the offspring are born with the agouti gene turned off[69] This effect has been used as an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome[70].

Most cytosine methylation occurs on cytosines located near guanines, called CpG sites. These CpG sites are often located upstream of promoters, or within the gene body. CpG islands are defined as regions that are greater than 500 bp in length with greater than 55% GC and an expected/observed CpG ratio of > 0.65.

While cytosine methylation (5mC) is known as a silencing mark that represses genes, cytosine hydroxymethylation (5hmC) is shown to be an activating mark that promotes gene expression and is a proposed intermediate in the DNA demethylation pathway[1,4,6]. Similar to 5mC, 5hmC is involved during development, cancers, cell differentiation, and diseases[71].

5mC and/or 5hmC can be a diagnostic tool to help identify the effects of nutrition, carcinogens[72], and environmental factors in relation to diseases. The impact of these modifications on gene regulation depends on their locations within the genome. It is therefore important to determine the exact position of the modified bases.

65 Smith Z. D. and Meissner A. (2013) DNA methylation: roles in mammalian development. Nat Rev Genet 14: 204-220

66 Jullien P. E. and Berger F. (2010) DNA methylation reprogramming during plant sexual reproduction? Trends Genet 26: 394-399

67 Schmitz R. J., He Y., Valdes-Lopez O., Khan S. M., Joshi T., et al. (2013) Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res 23: 1663-1674

68 Koh K. P. and Rao A. (2013) DNA methylation and methylcytosine oxidation in cell fate decisions. Curr Opin Cell Biol 25: 152-161

69 Dolinoy D. C., Weidman J. R., Waterland R. A. and Jirtle R. L. (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114: 567-572

70 Dolinoy D. C. (2008) The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev 66 Suppl 1: S7-11, Dolinoy D. C. and Faulk C. (2012) Introduction: The use of animals models to advance epigenetic science. ILAR J 53: 227-231

71 Pfeifer G. P., Kadam S. and Jin S. G. (2013) 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin 6: 10

72 Thomson J. P., Lempiainen H., Hackett J. A., Nestor C. E., Muller A., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. Genome Biol 13: R93

Base	Sequence	BS Sequence	oxBS Sequence	TAB Sequence	RRBS Sequence
C	C	T	T	T	T
5mC	C	C	C	T	C
5hmC	C	C	T	C	C

Sequencing reads created by various methods

Reviews

Koh K. P. and Rao A. (2013) DNA methylation and methylcytosine oxidation in cell fate decisions. Curr Opin Cell Biol 25: 152-161

Lister R., Mukamel E. A., Nery J. R., Urich M., Puddifoot C. A., et al. (2013) Global epigenomic reconfiguration during mammalian brain development. Science 341: 1237905

Pfeifer G. P., Kadam S. and Jin S. G. (2013) 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin 6: 10

Piccolo F. M. and Fisher A. G. (2014) Getting rid of DNA methylation. Trends Cell Biol 24: 136-143

Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

Schweiger M. R., Barmeyer C. and Timmermann B. (2013) Genomics and epigenomics: new promises of personalized medicine for cancer patients. Brief Funct Genomics 12: 411-421

Smith Z. D. and Meissner A. (2013) DNA methylation: roles in mammalian development. Nat Rev Genet 14: 204-220

Telese F., Gamliel A., Skowronska-Krawczyk D., Garcia-Bassets I. and Rosenfeld M. G. (2013) “Seq-ing” insights into the epigenetics of neuronal gene regulation. Neuron 77: 606-623

Veluchamy A., Lin X., Maumus F., Rivarola M., Bhavsar J., et al. (2013) Insights into the role of DNA methylation in diatoms by genome-wide profiling in Phaeodactylum tricornutum. Nat Commun 4: 2091

Vidaki A., Daniel B. and Court D. S. (2013) Forensic DNA methylation profiling--Potential opportunities and challenges. Forensic Sci Int Genet 7: 499-507

References

Meaburn E. and Schulz R. (2012) Next generation sequencing in epigenetics: insights and challenges. Semin Cell Dev Biol 23: 192-199

Thomson J. P., Lempiainen H., Hackett J. A., Nestor C. E., Muller A., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. Genome Biol 13: R93

Jin S. G., Kadam S. and Pfeifer G. P. (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res 38: e125

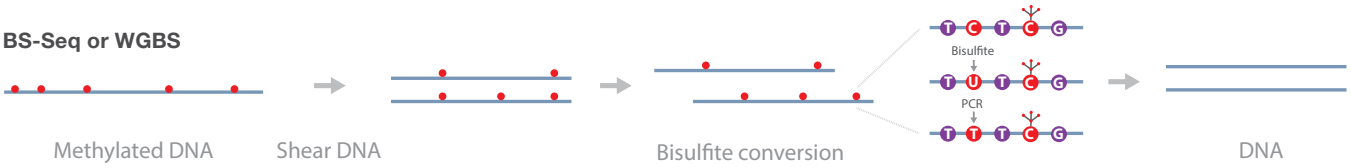
Dolinoy D. C., Weidman J. R., Waterland R. A. and Jirtle R. L. (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114: 567-572

BISULFITE SEQUENCING (BS-SEQ)

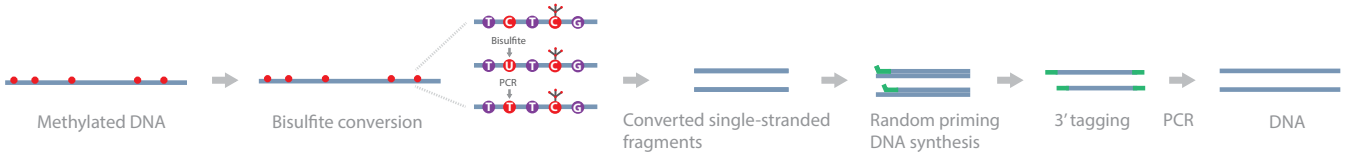
Bisulfite sequencing (BS-Seq) or whole-genome bisulfite sequencing (WGBS) is a well-established protocol to detect methylated cytosines in genomic DNA[73]. In this method, genomic DNA is treated with sodium bisulfite and then sequenced, providing single-base resolution of methylated cytosines in the genome. Upon bisulfite treatment, unmethylated cytosines are deaminated to uracils which, upon sequencing, are converted to thymidines. Simultaneously, methylated cytosines resist deamination and are read as cytosines. The location of the methylated cytosines can then be determined by comparing treated and untreated sequences. Bisulfite treatment of DNA converts unmethylated cytosines to thymidines, leading to reduced sequence complexity. Very accurate deep sequencing serves to mitigate this loss of complexity

The EpiGnome™ Kit uses a unique library construction method that incorporates bisulfite conversion as the first step. The EpiGnome method retains sample diversity while providing uniform coverage.

BS-Seq or WGBS



EpiGnome Methyl-Seq



Pros	Cons
BS-Seq or WGBS	
• *CpG and non-CpG methylation throughout the genome is covered at single-base resolution* • *5mC in dense, less dense, and repeat regions are covered*	• *Bisulfite converts unmethylated cytosines to thymidines, reducing sequence complexity, which can make it difficult to create alignments* • *NPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion* • *Bisulfite conversion does not distinguish between 5mC and 5hmC*

EpiGnome	
• *Pre-library bisulfite conversion* • *Low input gDNA (50 ng)* • *Uniform CpG, CHG, and CHH coverage* • *No fragmentation and no methylated adapters* • *Retention of sample diversity*	• *Bisulfite converts unmethylated cytosines to thymidines,reducing sequence complexity, which can make it difficult to create alignments* • *SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion* • *Bisulfite conversion does not distinguish between 5mC and 5hmC* • *Higher duplicate percentage*

73 Feil R., Charlton J., Bird A. P., Walter J. and Reik W. (1994) Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. Nucleic Acids Res 22: 695-696

References

Gustems M., Woellmer A., Rothbauer U., Eck S. H., Wieland T., et al. (2013) c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. Nucleic Acids Res

Transcription factors bind with specificity to their preferred DNA sequence motif. However, a virus-encoded transcription factor Zta was the first example of a sequence-specific transcription factor binding selectively and preferentially to methylated CpG residues. In this study the authors present their finding of a novel AP-1 binding site, termed meAP-1, which contains a CpG nucleotide. Using ChIP-Seq with Illumina sequencing, they show how the methylation state of this nucleotide affects binding by c-Jun/c-Fos in vitro and in vivo.
Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Habibi E., Brinkman A. B., Arand J., Kroeze L. I., Kerstens H. H., et al. (2013) Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell 13: 360-369

Mouse embryonic stem cells (ESCs) provide an excellent model system for studying mammalian cell differentiation on the molecular level. This study uses two kinase inhibitors (2i) to derive mouse ESCs in the pluripotent ground state to study the deposition and loss of DNA methylation during differentiation. The epigenetic state and expression of the cells were monitored using ChIP-Seq and RNA-Seq on the Illumina HiSeq platform.

Illumina Technology: HiSeq 2000, MiSeq

Hussain S., Sajini A. A., Blanco S., Dietmann S., Lombard P., et al. (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4: 255-261

This paper presents miCLIP: a new technique for identifying RNA methylation sites in transcriptomes. The authors use the miCLIP method with Illumina sequencing to determine site-specific methylation in tRNAs and additional messenger and noncoding RNAs. As a case study, the authors studied the methyltransferase NSun2 and showed that loss of cytosine-5 methylation in vault RNAscauses aberrant processing that may interrupt processing of small RNA fragments, such as microRNAs.

Illumina Technology: TruSeq RNA Kit, Genome Analyzer$_{IIx}$

Kozlenkov A., Roussos P., Timashpolsky A., Barbu M., Rudchenko S., et al. (2014) Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. Nucleic Acids Res 42: 109-127

Epigenetic regulation by DNA methylation varies among different cell types. In this study, the authors compared the methylation status of neuronal and non-neuronal nuclei using Illumina Human Methylation450k arrays. They classified the differentially methylated (DM) sites into those undermethylated in the neuronal cell type, and those that were undermethylated in non-neuronal cells. Using this approach, they identified sets of cell type–specific patterns and characterized these by their genomic locations.

Illumina Technology: HumanMethylation450 BeadChip, HumanOmni1-Quad (Infinium GT), HiSeq 2000

Lun F. M., Chiu R. W., Sun K., Leung T. Y., Jiang P., et al. (2013) Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. Clin Chem 59: 1583-1594

The presence of fetal DNA in maternal plasma opens up possibilities for non-invasive prenatal DNA testing of the fetus through blood samples from the mother. Using SNP differences between mother and fetus to identify fetal molecules, this study inspected the genome-wide methylome of the unborn child by bisulfite sequencing. The authors determined the methylation density over each 1 Mbp region of the genome for samples taken in each trimester and after delivery to show how the fetal methylome is established gradually throughout pregnancy.

Illumina Technology: HiSeq 2000, HumanMethylation450 BeadChip

Regulski M., Lu Z., Kendall J., Donoghue M. T., Reinders J., et al. (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res 23: 1651-1662

The maize genome encompasses a widely unexplored landscape for epigenetic mechanisms of paramutation and imprinting. In this study whole-exome bisulfite sequencing was applied to map the cytosine methylation profile of two maize inbred lines. The analysis revealed that frequent methylation switches, guided by siRNA, may persist for up to eight generations, suggesting that epigenetic inheritance resembling paramutation is much more common than previously supposed.

Illumina Technology: HiSeq 2000, Genome Analyzer$_{IIx}$

Schmitz R. J., He Y., Valdes-Lopez O., Khan S. M., Joshi T., et al. (2013) Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res 23: 1663-1674

In an effort to elucidate the mammalian DNA methylome, this study applied whole-genome bisulfite sequencing using the Illumina HiSeq platform and gene expression analysis to define functional classes of hypomethylated regions (HMRs). Comparing HMR profiles in embryonic stem and primary blood cells, the authors showed that the HMRs in intergenic space (iHMRs) mark an exclusive subset of active DNase hypersensitive sites. The authors went on to compare primate-specific and human population variation at iHMRs, and they derived models of the cellular timelines for DHS and iHMR establishment.

Illumina Technology: HiSeq 2000

Schlesinger F., Smith A. D., Gingeras T. R., Hannon G. J. and Hodges E. (2013) De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements. Genome Res 23: 1601-1614

In an effort to elucidate the mammalian DNA methylome, this study applied whole-genome bisulfite sequencing using the Illumina HiSeq platform and gene expression analysis to define functional classes of hypomethylated regions (HMRs). Comparing HMR profiles in embryonic stem and primary blood cells, the authors showed that the HMRs in intergenic space (iHMRs) mark an exclusive subset of active DNase hypersensitive sites. The authors went on to compare primate-specific and human population variation at iHMRs, and they derived models of the cellular timelines for DHS and iHMR establishment.

Illumina Technology: HiSeq 2000

Xie W., Schultz M. D., Lister R., Hou Z., Rajagopal N., et al. (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 153: 1134-1148

The authors studied the differentiation of hESCs into four cell types: trophoblast-like cells, mesendoderm, neural progenitor cells, and mesenchymal stem cells. DNA methylation (WGBS) and histone modifications were examined for each cell type. The study provides insight into the dynamic changes that accompany lineage-specific cell differentiation in hESCs.

Illumina Technology: HiSeq 2000

Yamaguchi S., Shen L., Liu Y., Sendler D. and Zhang Y. (2013) Role of Tet1 in erasure of genomic imprinting. Nature 504: 460-464

Genomic imprinting is the cellular mechanism for switching off one of two alleles by DNA methylation. This allele-specific gene expression system is very important for mammalian development and function. In this study, the Tet1 protein was studied for its function in primordial germ cells, the phase of development where the imprinting methylation mark of the parent is erased. Using ChIP-Seq and bisulfite sequencing on the Illumina HiSeq platform, the authors showed that Tet1 knockout males exhibited aberrant hypermethylation in the paternal allele of differential methylated regions.

Illumina Technology: HiSeq 2500®

Blaschke K., Ebata K. T., Karimi M. M., Zepeda-Martinez J. A., Goyal P., et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500: 222-226

Potok M. E., Nix D. A., Parnell T. J. and Cairns B. R. (2013) Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. Cell 153: 759-772

Rodrigues J. A., Ruan R., Nishimura T., Sharma M. K., Sharma R., et al. (2013) Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. Proc Natl Acad Sci U S A 110: 7934-7939

Shirane K., Toh H., Kobayashi H., Miura F., Chiba H., et al. (2013) Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet 9: e1003439

Warden C. D., Lee H., Tompkins J. D., Li X., Wang C., et al. (2013) COHCAP: an integrative genomic pipeline for single-nucleotide resolution DNA methylation analysis. Nucleic Acids Res 41: e117

Adey A. and Shendure J. (2012) Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. Genome Res 22: 1139-1143

Diep D., Plongthongkum N., Gore A., Fung H. L., Shoemaker R., et al. (2012) Library-free methylation sequencing with bisulfite padlock probes. Nat Methods 9: 270-272

Seisenberger S., Andrews S., Krueger F., Arand J., Walter J., et al. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell 48: 849-862

Feng S., Cokus S. J., Zhang X., Chen P. Y., Bostick M., et al. (2010) Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A 107: 8689-8694

Li N., Ye M., Li Y., Yan Z., Butcher L. M., et al. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. Methods 52: 203-212

Lyko F., Foret S., Kucharski R., Wolf S., Falckenhayn C., et al. (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol 8: e1000506

Ball M. P., Li J. B., Gao Y., Lee J. H., LeProust E. M., et al. (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 27: 361-368

Gehring M., Bubb K. L. and Henikoff S. (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324: 1447-1451

Hodges E., Smith A. D., Kendall J., Xuan Z., Ravi K., et al. (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res 19: 1593-1605

Hsieh T. F., Ibarra C. A., Silva P., Zemach A., Eshed-Williams L., et al. (2009) Genome-wide demethylation of Arabidopsis endosperm. Science 324: 1451-1454

Jacob Y., Feng S., Leblanc C. A., Bernatavichute Y. V., Stroud H., et al. (2009) ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat Struct Mol Biol 16: 763-768

Cokus S. J., Feng S., Zhang X., Chen Z., Merriman B., et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452: 215-219

He Y., Vogelstein B., Velculescu V. E., Papadopoulos N. and Kinzler K. W. (2008) The antisense transcriptomes of human cells. Science 322: 1855-1857

Meissner A., Mikkelsen T. S., Gu H., Wernig M., Hanna J., et al. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454: 766-770

Associated Kits

EpiGnome™ Methyl-Seq® Kit

Infinium HumanMethylation450® Arrays

POST-BISULFITE ADAPTER TAGGING (PBAT)

To avoid the bisulfite-induced loss of intact sequencing templates, in post-bisulfite adapter tagging (PBAT)[74] bisulfite treatment is followed by adapter tagging and two rounds of random primer extension. This procedure generates a substantial number of unamplified reads from as little as subnanogram quantities of DNA.

Pros	Cons
• *Requires only 100 ng of DNA for amplification-free WGBS of mammalian genomes*	• *Bisulfite converts unmethylated cytosines to thymidines, reducing sequence complexity, which can make it difficult to create alignments* • *SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion* • *Bisulfite conversion does not distinguish between 5mC and 5hmC*

References

Kobayashi H., Sakurai T., Miura F., Imai M., Mochiduki K., et al. (2013) High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. Genome Res 23: 616-627

Dynamic epigenetic reprogramming occurs during mammalian germ cell development. One of these processes is DNA methylation and demethylation, which is commonly studied using bisulfite sequencing. This study used an Illumina HiSeq 2000 system for WGBS to characterize the DNA methylation profiles of male and female mouse primordial germ cells (PGCs) at different stages of embryonic development. The authors found sex- and chromosome-specific differences in genome-wide CpG and CGI methylation during early- to late-stage PGC development. They also obtained high-resolution details of DNA methylation changes, for instance, that LINE/LTR retrotransposons were resistant to DNA methylation at high CpG densities.

Illumina Technology: HiSeq 2000

74 Miura F., Enomoto Y., Dairiki R. and Ito T. (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res 40: e136

Shirane K., Toh H., Kobayashi H., Miura F., Chiba H., et al. (2013) Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet 9: e1003439

DNA methylation is an epigenetic modification that plays a crucial role in normal mammalian development, retrotransposon silencing, and cellular reprogramming. Using amplification-free WGBS, the authors constructed the base-resolution methylome maps of germinal vesicle oocytes (GVOs), non-growing oocytes (NGOs), and mutant GVOs lacking the DNA methyltransferases Dnmt1, Dnmt3a, Dnmt3b, or Dnmt3L. They found that nearly two-thirds of all methylcytosines occur in a non-CG context in GVOs. The distribution of non-CG methylation closely resembled that of CG methylation throughout the genome and showed clear enrichment in gene bodies.

Illumina Technology: HiSeq 2000

Kobayashi H., Sakurai T., Imai M., Takahashi N., Fukuda A., et al. (2012) Contribution of Intragenic DNA Methylation in Mouse Gametic DNA Methylomes to Establish Oocyte-Specific Heritable Marks. PLoS Genet 8: e1002440

Miura F., Enomoto Y., Dairiki R. and Ito T. (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res 40: e136

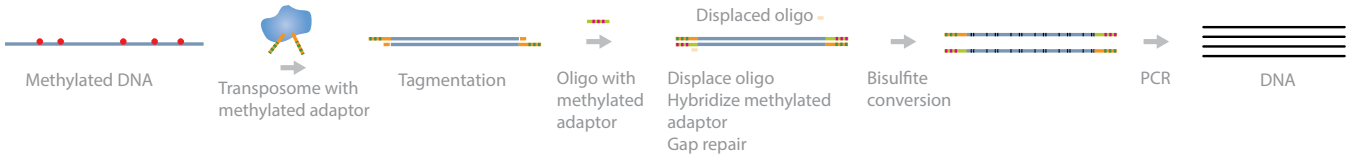
Associated Kits

EpiGnome™ Methyl-Seq Kit

Infinium HumanMethylation450 Arrays

TAGMENTATION-BASED WHOLE GENOME BISULFITE SEQUENCING (T-WGBS)

Tagmentation-based whole-genome bisulfite sequencing (T-WGBS) is a protocol that utilizes the Epicentre® Tn5 transposome and bisulfite conversion to study 5mC[75]. In this method, DNA is incubated with Tn5 transposome containing methylated primers, which fragments the DNA and ligates adapters. Tagged DNA first undergoes oligo displacement, followed by methylated oligo replacement and gap repair, assuring methylated adapter addition to tagmented DNA. DNA is then treated with sodium bisulfite, PCR-amplified, and sequenced. Deep sequencing provides single-base resolution of 5mC in the genome.



Pros	Cons
• *Can sequence samples with very limited starting material (~20 ng)* • *Fast protocol with few steps* • *Elimination of multiple steps prevents loss of DNA*	• *Bisulfite converts unmethylated cytosines to thymidines, reducing sequence complexity, which can make it difficult to create alignments* • *SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion* • *Bisulfite conversion does not distinguish between 5mC and 5hmC*

References

Wang Q., Gu L., Adey A., Radlwimmer B., Wang W., et al. (2013) Tagmentation-based whole-genome bisulfite sequencing. Nat Protoc 8: 2022-2032

Scaling up bisulfite sequencing to genome-wide analysis has been hindered by the requirements for large amounts of DNA and high sequencing costs. This paper presents a protocol for T-WGBS with sequencing on the Illumina HiSeq 2000 system. The authors demonstrate the robustness of the protocol in comparison with conventional WGBS. T-WGBS requires not more than 20 ng of input DNA; hence, the protocol allows the comprehensive methylome analysis of limited amounts of DNA isolated from precious biological specimens.

Illumina Technology: Nextera DNA Sample Prep Kit, HiSeq 2000; 101 bp paired-end reads

Associated Kits

EpiGnome™ Methyl-Seq Kit

Infinium HumanMethylation450 Arrays

Nextera DNA Sample Prep Kit

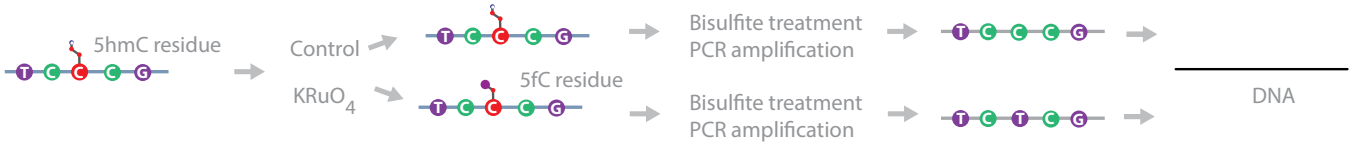
Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

75 Wang Q., Gu L., Adey A., Radlwimmer B., Wang W., et al. (2013) Tagmentation-based whole-genome bisulfite sequencing. Nat Protoc 8: 2022-2032

OXIDATIVE BISULFITE SEQUENCING (OXBS-SEQ)

Oxidative bisulfite sequencing (oxBS-Seq) differentiates between 5mC and 5hmC[76]. With oxBS, 5hmC is oxidized to 5formylcytosine (5fC) with an oxidizing agent, while 5mC remains unchanged. Sodium bisulfite treatment of oxidized 5hmC results in its deamination to uracil which, upon sequencing, is read as a thymidine. Deep sequencing of oxBS-treated DNA and sequence comparison of treated vs. untreated can identify 5mC locations at base resolution.



Pros	Cons
• *CpG and non-CpG methylation throughout the genome is covered at single-base resolution* • *5mC dense and less dense in repeat regions are covered* • *Method clearly differentiates between 5mC and 5hmC, precisely identifying 5mC*	• *Bisulfite converts unmethylated cytosines to thymidines,reducing sequence complexity, which can make it difficult to create alignments* • *SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion*

References

Booth M. J., Ost T. W., Beraldi D., Bell N. M., Branco M. R., et al. (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat Protoc 8: 1841-1851

This is the first paper to report a method combining chemical treatment of DNA with the well-established bisulfite protocol, highlighting Illumina's TruSeq kit and calling for the use of MiSeq or HiSeq platforms. The OxBS-Seq protocol helps distinguish between 5mC and 5hmC, while standard bisulfite sequencing is incapable of distinguishing between 5mC and 5hmC. Genomic DNA is first treated with an oxidizing agent that reacts with 5hmC, promoting its deamination to uracil, while the 5mC modification remains unchanged and is read as cytosine. Using Illumina technology, this method allows base resolution of the exact location of 5hmC and 5mC modifications.

Illumina Technology: TruSeq DNA Sample Prep Kit, MiSeq, HiSeq 2000

Associated Kits

EpiGnome™ Methyl-Seq Kit

TruSeq DNA Sample Prep Kit

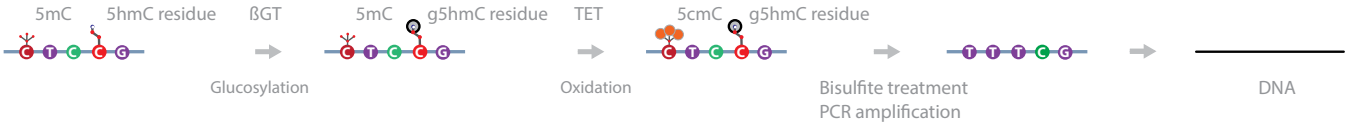
TruSeq DNA PCR-Free Sample Prep Kit

TruSeq Nano DNA Sample Prep Ki

76 Booth M. J., Branco M. R., Ficz G., Oxley D., Krueger F., et al. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 336: 934-937

TET-ASSISTED BISULFITE SEQUENCING (TAB-SEQ)

TAB-Seq is a novel method that uses bisulfite conversion and Tet proteins to study 5hmC[77]. In this protocol, 5hmC is first protected with a glucose moiety that allows selective interaction and subsequent oxidation of 5mC with the Tet proteins. The oxidized genomic DNA is then treated with bisulfite, where 5hmC remains unchanged and is read as a cytosine, while 5mC and unmethylated cytosines are deaminated to uracil and read as thymidines upon sequencing. Deep sequencing of TAB-treated DNA compared with untreated DNA provides accurate representation of 5hmC localization in the genome.



Pros	Cons
• *CpG and non-CpG hydroxymethylation throughout the genome is covered at single-base resolution* • *Dense, less dense, and 5hmC in repeat regions are covered* • *Method clearly differentiates between 5hmC and 5mC, specifically identifying 5hmC*	• *Bisulfite converts unmethylated cytosines to thymidines,reducing sequence complexity, which can make it difficult to create alignments* • *SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion* • *Requires deep sequencing to provide sufficient depth to cover the entire genome and accurately map the low amounts 5hmC*[78]

References

Kim M., Park Y. K., Kang T. W., Lee S. H., Rhee Y. H., et al. (2013) Dynamic changes in DNA methylation and hydroxymethylation when hES cells undergo differentiation toward a neuronal lineage. Hum Mol Genet 23: 657-667

Epigenetic markers on chromatin include the methylation of DNA. Several forms of DNA methylation exist and their function and interaction is a field of intensive study. This paper describes how an in vitro model system of gradual differentiation of hESCs underwent dramatic genome-wide changes in 5mC and 5hmC methylationpatterns during lineage commitment. The authors used Illumina BeadArray for expression profiling and Genome Analyzer hMeDIP-sequencing to study the correlation between gene expression and DNA methylation.

Illumina Technology: Human-6 Whole-Genome Expression BeadChip, Genome Analyzer$_{IIx}$, HiScanSQ® Scanner, Infinium HumanMethylation 450 BeadChip

77 Yu M., Hon G. C., Szulwach K. E., Song C. X., Zhang L., et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the Mammalian genome. Cell 149: 1368-1380

78 Thomson J. P., Hunter J. M., Nestor C. E., Dunican D. S., Terranova R., et al. (2013) Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. Nucleic Acids Res 41: e206

Lister R., Mukamel E. A., Nery J. R., Urich M., Puddifoot C. A., et al. (2013) Global epigenomic reconfiguration during mammalian brain development. Science 341: 1237905

DNA methylation is implicated in mammalian brain development and plasticity underlying learning and memory. This paper reports the genome-wide composition, patterning, cell specificity, and dynamics of DNA methylation at single-base resolution in human and mouse frontal cortex throughout their lifespan. The extensive methylome profiling was performed with ChIP-Seq on an Illumina HiSeq sequencer at single-base resolution.

Illumina Technology: TruSeq RNA Sample Prep Kit, TruSeq DNA Sample Prep Kit, HiSeq 2000

Wang T., Wu H., Li Y., Szulwach K. E., Lin L., et al. (2013) Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency. Nat Cell Biol 15: 700-711

The transcriptional reprogramming that allows mammalian somatic cells to be reprogrammed into pluripotent stem cells (iPSCs) includes a complete reconfiguration of the epigenetic marks in the genome. This study examined the levels of 5hmC in hESCs during reprogramming to iPSCs. The authors found reprogramming hotspots in subtelomeric regions, most of which featured incomplete hydroxymethylation at CG sites.

Illumina Technology: HiSeq 2000, HiScanSQ, MiSeq

Jiang L., Zhang J., Wang J. J., Wang L., Zhang L., et al. (2013) Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. Cell 153: 773-784

Song C. X., Szulwach K. E., Dai Q., Fu Y., Mao S. Q., et al. (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell 153: 678-691

Yu M., Hon G. C., Szulwach K. E., Song C. X., Jin P., et al. (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. Nat Protoc 7: 2159-2170

Yu M., Hon G. C., Szulwach K. E., Song C. X., Zhang L., et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 149: 1368-1380

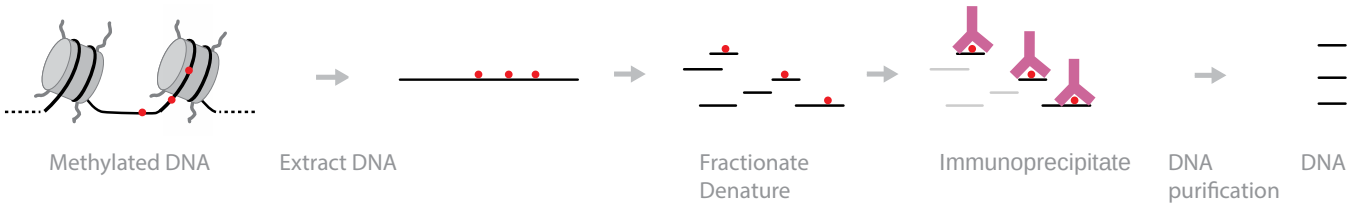
Associated Kits

EpiGnome™ Methyl-Seq Kit

Infinium HumanMethylation450 Arrays

METHYLATED DNA IMMUNOPRECIPITATION SEQUENCING (MEDIP-SEQ)

Methylated DNA immunoprecipitation sequencing (MeDIP-Seq) is commonly used to study 5mC or 5hmC modification[79] . Specific antibodies can be used to study cytosine modifications. If using 5mC-specific antibodies, methylated DNA is isolated from genomic DNA via immunoprecipitation. Anti-5mC antibodies are incubated with fragmented genomic DNA and precipitated, followed by DNA purification and sequencing. Deep sequencing provides greater genome coverage, representing the majority of immunoprecipitated methylated DNA.



Pros	Cons
• *Covers CpG and non-CpG 5mC throughout the genome* • *5mC in dense, less dense, and repeat regions are covered* • *Antibody-based selection is independent of sequence and does not enrich for 5hmC due to antibody specificity*	• *Base-pair resolution is lower (~150 bp) as opposed to single base resolution* • *Antibody specificity and selectivity must be tested to avoid nonspecific interaction* • *Antibody-based selection is biased towards hypermethylated regions*

References

Puszyk W., Down T., Grimwade D., Chomienne C., Oakey R. J., et al. (2013) The epigenetic regulator PLZF represses L1 retrotransposition in germ and progenitor cells. EMBO J 32: 1941-1952

Each transcription factor in the human cell may regulate a large number of target genes through specific chromatin interactions. Promyelocytic leukemia zinc finger protein (PLZF) acts as an epigenetic regulator of stem cell maintenance in germ cells and hematopoietic stem cells. In this study, L1 retrotransposons were identified as the primary targets of PLZF. Using ChIP-Seq and MeDIP-Seq onIllumina Genome Analyzer, the authors identified how PLZF-mediated DNA methylation induces silencing of L1 and inhibits L1 retrotransposition.

Illumina Technology: Genome Analyzer$_{IIx}$

Shen H., Qiu C., Li J., Tian Q. and Deng H. W. (2013) Characterization of the DNA methylome and its interindividual variation in human peripheral blood monocytes. Epigenomics 5: 255-269

Peripheral blood monocytes (PBMs) play multiple and critical roles in the immune response, and abnormalities in PBMs have been linked to a variety of human disorders. In this study, the epigenome-wide DNA methylation profiles of purified PBMs were identified using MeDIP-Seq on an Illumina Genome Analyzer. Interestingly, the authors observed substantial interindividual variation in DNA methylation across the individual PBM methylomes.

Illumina Technology: Genome Analyzer$_{IIx}$

79 Weber M., Davies J. J., Wittig D., Oakeley E. J., Haase M., et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37: 853-862

Tan L., Xiong L., Xu W., Wu F., Huang N., et al. (2013) Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. Nucleic Acids Res 41: e84

The genome-wide distribution patterns of the “sixth base” 5hmC in many tissues and cells have recently been revealed by hydroxymethylated DNA immunoprecipitation (hMeDIP) followed by high throughput sequencing or tiling arrays. This paper presents a new comparative hMeDIP-seq method which allows for direct genome-wide comparison of DNA hydroxymethylation across multiplesamples. The authors demonstrate the new method by profiling DNA hydroxymethylation and gene expression during neural differentiation.

Illumina Technology: Genome Analyzer$_{IIx}$

Saied M. H., Marzec J., Khalid S., Smith P., Down T. A., et al. (2012) Genome wide analysis of acute myeloid leukemia reveal leukemia specific methylome and subtype specific hypomethylation of repeats. PLoS One 7: e33213

Epigenetic modifications in the form of DNA methylation are part of the regulatory machinery of the cell. By studying the patterns of DNA methylation in disease tissue, we may characterize disease mechanisms. In this study, bone marrow samples from 12 patients with acute myeloid leukemia (AML) were analyzed with MeDIP-Seq and compared to normal bone marrow. The investigators found considerable cytogenetic subtype specificity in the methylomes affecting different genomic features.

Illumina Technology: HumanMethylation27 arrays, Genome Analyzer$_{IIx}$

Taiwo O., Wilson G. A., Morris T., Seisenberger S., Reik W., et al. (2012) Methylome analysis using MeDIP-seq with low DNA concentrations. Nat Protoc 7: 617-636

DNA methylation can be assayed at high throughput using MeDIP-Seq, but the application has been limited to samples where the amount of DNA was sufficient for the assay (5–20 µg). This study presents a new optimized protocol for MeDIP-Seq, requiring as little as 50 ng of starting DNA.

Illumina Technology: Genome Analyzer$_{IIx}$

Bian C. and Yu X. (2013) PGC7 suppresses TET3 for protecting DNA methylation. Nucleic Acids Res

Colquitt B. M., Allen W. E., Barnea G. and Lomvardas S. (2013) Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc Natl Acad Sci U S A 110: 14682-14687

Neri F., Krepelova A., Incarnato D., Maldotti M., Parlato C., et al. (2013) Dnmt3L Antagonizes DNA Methylation at Bivalent Promoters and Favors DNA Methylation at Gene Bodies in ESCs. Cell 155: 121-134

Stevens M., Cheng J. B., Li D., Xie M., Hong C., et al. (2013) Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. Genome Res 23: 1541-1553

Zhang B., Zhou Y., Lin N., Lowdon R. F., Hong C., et al. (2013) Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Res 23: 1522-1540

Zilbauer M., Rayner T. F., Clark C., Coffey A. J., Joyce C. J., et al. (2013) Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type-specific hypomethylated regions. Blood 122: e52-60

Sati S., Tanwar V. S., Kumar K. A., Patowary A., Jain V., et al. (2012) High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. PLoS One 7: e31621

Gao Q., Steine E. J., Barrasa M. I., Hockemeyer D., Pawlak M., et al. (2011) Deletion of the de novo DNA methyltransferase Dnmt3a promotes lung tumor progression. Proc Natl Acad Sci U S A 108: 18061-18066

Bock C., Tomazou E. M., Brinkman A. B., Muller F., Simmer F., et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114

Chavez L., Jozefczuk J., Grimm C., Dietrich J., Timmermann B., et al. (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. Genome Res 20: 1441-1450

Harris R. A., Wang T., Coarfa C., Nagarajan R. P., Hong C., et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol 28: 1097-1105

Li N., Ye M., Li Y., Yan Z., Butcher L. M., et al. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. Methods 52: 203-212

Maunakea A. K., Nagarajan R. P., Bilenky M., Ballinger T. J., D'Souza C., et al. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466: 253-257

Ruike Y., Imanaka Y., Sato F., Shimizu K. and Tsujimoto G. (2010) Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. BMC Genomics 11: 137

Hammoud S. S., Nix D. A., Zhang H., Purwar J., Carrell D. T., et al. (2009) Distinctive chromatin in human sperm packages genes for embryo development. Nature 460: 473-478

Pomraning K. R., Smith K. M. and Freitag M. (2009) Genome-wide high throughput analysis of DNA methylation in eukaryotes. Methods 47: 142-150

Down T. A., Rakyan V. K., Turner D. J., Flicek P., Li H., et al. (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 26: 779-785

Associated Kits

Infinium HumanMethylation450 Arrays

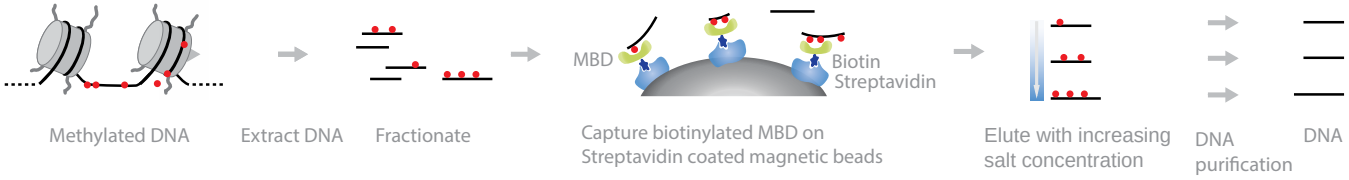
Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

METHYLATION-CAPTURE (METHYLCAP) SEQUENCING OR METHYL-BINDING-DOMAIN–CAPTURE (MBDCAP) SEQUENCING

MethylCap[80, 81] or MBDCap[82, 83] uses proteins to capture methylated DNA in the genome. Genomic DNA is first sonicated and incubated with tagged MBD proteins that can bind methylated cytosines. The protein-DNA complex is then precipitated with antibody-conjugated beads that are specific to the protein tag. Deep sequencing provides greater genome coverage, representing the majority of MBD-bound methylated DNA.



Pros	Cons
• *Genome-wide coverage of 5mC in dense CpG areas and repeat regions* • *MBD proteins do not interact with 5hmC*	• *Genome-wide CpGs and non-CpG methylation is not covered Areas with less dense 5mC are also missed* • *Base-pair resolution is lower (~150 bp) as opposed to single base resolution* • *Protein-based selection is biased towards hypermethylated regions*

References

Kim M., Park Y. K., Kang T. W., Lee S. H., Rhee Y. H., et al. (2013) Dynamic changes in DNA methylation and hydroxymethylation when hES cells undergo differentiation toward a neuronal lineage. Hum Mol Genet 23: 657-667

Epigenetic markers on chromatin include the methylation of DNA. Several forms of DNA methylation exist and their function and interaction is a field of intensive study. This paper describes how an in vitro model system of gradual differentiation of hESCs underwent dramatic genome-wide changes in 5mC and 5hmC methylationpatterns during lineage commitment. The authors used Illumina BeadArray for expression profiling and Genome Analyzer hMeDIP-sequencing to study the correlation between gene expression and DNA methylation.

Illumina Technology: Human-6 Whole-Genome Expression BeadChip, Genome Analyzer$_{IIx}$, HiScanSQ Scanner, Infinium HumanMethylation 450 BeadChip

80 Bock C., Tomazou E. M., Brinkman A. B., Muller F., Simmer F., et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114

81 Brinkman A. B., Simmer F., Ma K., Kaan A., Zhu J., et al. (2010) Whole-genome DNA methylation profiling using MethylCap-seq. Methods 52: 232-236

82 Rauch T. A., Zhong X., Wu X., Wang M., Kernstine K. H., et al. (2008) High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U S A 105: 252-257

83 Rauch T. A. and Pfeifer G. P. (2009) The MIRA method for DNA methylation analysis. Methods Mol Biol 507: 65-75

Huang T. T., Gonzales C. B., Gu F., Hsu Y. T., Jadhav R. R., et al. (2013) Epigenetic deregulation of the anaplastic lymphoma kinase gene modulates mesenchymal characteristics of oral squamous cell carcinomas. Carcinogenesis 34: 1717-1727

Promoter methylation is associated with silencing tumor suppressor genes in oral squamous cell carcinomas (OSCCs). The authors used MBDCap-Seq to study methylation in OSCC cell lines, sequencing on the Illumina HiSeq platform, and identifying differentially methylated regions. The authors note the ALK gene was susceptible to epigenetic silencing during oral tumorigenesis.

Illumina Technology: HiSeq 2000

Zhao Y., Guo S., Sun J., Huang Z., Zhu T., et al. (2012) Methylcap-seq reveals novel DNA methylation markers for the diagnosis and recurrence prediction of bladder cancer in a Chinese population. PLoS ONE 7: e35175

Bladder cancer (BC) has a high mortality rate and is the sixth most common cancer in the world. For successfully treated BCs, the relapse rate is 60-70% within the first 5 years, necessitating the development of efficient diagnostics and biomarkers for monitoring disease progression. The presence of cells in the urine allow for noninvasive genetic screening directly from urine. In this study, the authors identify and validate nine DNA methylation markers through genome-wide profiling of DNA methylation from clinical urine samples.

Illumina Technology: Genome Analyzer$_{IIx}$

Brinkman A. B., Gu H., Bartels S. J., Zhang Y., Matarese F., et al. (2012) Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 22: 1128-1138

Rodriguez B. A., Frankhouser D., Murphy M., Trimarchi M., Tam H. H., et al. (2012) Methods for high-throughput MethylCap-Seq data analysis. BMC Genomics 13 Suppl 6: S14

Yu W., Jin C., Lou X., Han X., Li L., et al. (2011) Global analysis of DNA methylation by Methyl-Capture sequencing reveals epigenetic control of cisplatin resistance in ovarian cancer cell. PLoS One 6: e29450

Bock C., Tomazou E. M., Brinkman A. B., Muller F., Simmer F., et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114

Brinkman A. B., Simmer F., Ma K., Kaan A., Zhu J., et al. (2010) Whole-genome DNA methylation profiling using MethylCap-seq. Methods 52: 232-236

Serre D., Lee B. H. and Ting A. H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res 38: 391-399

Associated Kits

Infinium HumanMethylation450 Arrays

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Rapid Capture Exome/Custom Enrichment

REDUCED-REPRESENTATION BISULFITE SEQUENCING (RRBS-SEQ)

Reduced-representation bisulfite sequencing (RRBS-Seq) is a protocol that uses one or multiple restriction enzymes on the genomic DNA to produce sequence-specific fragmentation[84]. The fragmented genomic DNA is then treated with bisulfite and sequenced. This is the method of choice to study specific regions of interest. It is particularly effective where methylation is high, such as in promoters and repeat regions.

Pros	Cons
• *Genome-wide coverage of CpGs in islands at single-base resolution* • *Areas dense in CpG methylation are covered*	• *Restriction enzymes cut at specific sites, providing biased sequence selection* • *Method measures 10-15% of all CpGs in genome* • *Cannot distinguish between 5mC and 5hmC* • *Does not cover non-CpG areas, genome-wide CpGs, and CpGs in areas without the enzyme restriction site*

References

Kozlenkov A., Roussos P., Timashpolsky A., Barbu M., Rudchenko S., et al. (2014) Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. Nucleic Acids Res 42: 109-127

Epigenetic regulation by DNA methylation varies among different cell types. In this study, the authors compared the methylation status of neuronal and non-neuronal nuclei using Illumina Human Methylation450k arrays. They classified the differentially methylated (DM) sites into those undermethylated in the neuronal cell type, and those that were undermethylated in non-neuronal cells. Using this approach, they identified sets of cell type–specific patterns and characterized these by their genomic locations.

Illumina Technology: HumanMethylation450 BeadChip, HumanOmni1-Quad (Infinium GT), HiSeq 2000

Schillebeeckx M., Schrade A., Lobs A. K., Pihlajoki M., Wilson D. B., et al. (2013) Laser capture microdissection-reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse. Nucleic Acids Res 41: e116

DNA methylation profiling by sequencing is challenging due to inaccurate cell enrichment methods and low DNA yields. This proof-of-concept study presents a new method for genome-wide DNA methylation profiling using down to 1 ng of input DNA. The method—laser-capture microdissection reduced-representation bisulfite sequencing (LCM-RRBS)—combines Illumina HiSeq sequencing with customized methylated adapter sequences and bisulfite-PCR. The protocol allows for base-pair resolution of methylated sites.

Illumina Technology: HiSeq 2000, MiSeq

84 Meissner A., Gnirke A., Bell G. W., Ramsahoye B., Lander E. S., et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33: 5868-5877

Stevens M., Cheng J. B., Li D., Xie M., Hong C., et al. (2013) Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. Genome Res 23: 1541-1553

Current methods for sequencing-based DNA methylation profiling are continuously improving, but each common method, on its own, is insufficient in providing a genome-wide single-CpG resolution of DNA methylation at a low cost. In this paper the authors present a novel algorithm, methylCRF, which enables integration of data from MeDIP-Seq and MRE-Seq to provide single-CpG classification of methylation state. The method provides similar or higher accuracy than any array or sequencing method on its own. The authors demonstrate the algorithm on whole-genome bisulfite sequencing on Illumina HiSeq 2000 systems and Methylation450 arrays.

Illumina Technology: HumanMethylation450 BeadChip, HumanOmni1-Quad (Infinium GT), HiSeq 2000

Will B., Vogler T. O., Bartholdy B., Garrett-Bakelman F., Mayer J., et al. (2013) Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. Nat Immunol 14: 437-445

This study evaluated genome-wide DNA cytosine methylation by enhanced reduced-representation bisulfite sequencing (ERRBS). DNA was digested with MspI, then end-repaired and ligated to paired-end Illumina sequencing adapters. This was followed by size selection and bisulfite treatment, clean-up, and PCR prior to sequencing.

Illumina Technology: HiSeq 2000

Xi Y., Bock C., Muller F., Sun D., Meissner A., et al. (2012) RRBSMAP: a fast, accurate and user-friendly alignment tool for reduced representation bisulfite sequencing. Bioinformatics 28: 430-432

Bock C., Kiskinis E., Verstappen G., Gu H., Boulting G., et al. (2011) Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 144: 439-452

Gu H., Smith Z. D., Bock C., Boyle P., Gnirke A., et al. (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc 6: 468-481

Gertz J., Varley K. E., Reddy T. E., Bowling K. M., Pauli F., et al. (2011) Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genet 7: e1002228

Smallwood S. A., Tomizawa S., Krueger F., Ruf N., Carli N., et al. (2011) Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat Genet 43: 811-814

Ziller M. J., Muller F., Liao J., Zhang Y., Gu H., et al. (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genet 7: e1002389

Bock C., Tomazou E. M., Brinkman A. B., Muller F., Simmer F., et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114

Smith Z. D., Gu H., Bock C., Gnirke A. and Meissner A. (2009) High-throughput bisulfite sequencing in mammalian genomes. Methods 48: 226-232

Associated Kits

EpiGnome™ Methyl-Seq Kit

Infinium HumanMethylation450 Arrays

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

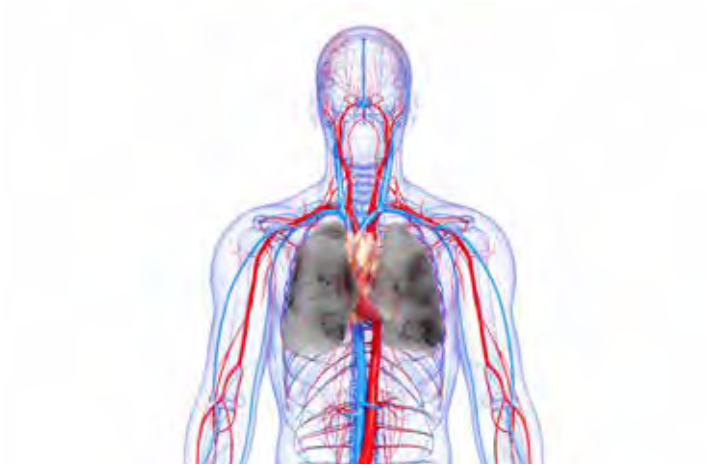
Nextera Rapid Capture Exome/Custom Enrichment Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

DNA-PROTEIN INTERACTIONS

Chromatin remodeling is a dynamic process driven by factors that change DNA-protein interactions. These epigenetic factors can involve protein modifications, such as histone methylation, acetylation, phosphorylation, and ubiquitination[85]. Histone modifications determine gene activation by recruiting regulatory factors and maintaining an open or closed chromatin state. Epigenetic factors play roles in tissue development[86], embryogenesis, cell fate, immune response, and diseases such as cancer[87]. Bacterial pathogens can elicit transcriptional repression of immune genes by chromatin remodeling[88]. The study of protein-DNA interactions has also demonstrated that chromatin remodeling can respond to external factors such as excessive alcohol-seeking behaviors[89], cigarette smoking[90], and clinical drugs.



Cigarette smoking disrupts DNA-protein interactions leading to the development of cancers or pulmonary diseases.

Reviews

Capell B. C. and Berger S. L. (2013) Genome-wide epigenetics. J Invest Dermatol 133: e9

Jakopovic M., Thomas A., Balasubramaniam S., Schrump D., Giaccone G., et al. (2013) Targeting the Epigenome in Lung Cancer: Expanding Approaches to Epigenetic Therapy. Front Oncol 3: 261

Kahrstrom C. T. (2013) Epigenetics: Legionella makes its mark on histones. Nat Rev Genet 14: 370 and local translation in neurons. Neuron 80: 648-657

Lee T. I. and Young R. A. (2013) Transcriptional regulation and its misregulation in disease. Cell 152: 1237-1251

Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

85 Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

86 Pinello L., Xu J., Orkin S. H. and Yuan G. C. (2014) Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. Proc Natl Acad Sci U S A 111: E344-353

87 Yeh H. H., Young D., Gelovani J. G., Robinson A., Davidson Y., et al. (2013) Histone deacetylase class II and acetylated core histone immunohistochemistry in human brains with Huntington's disease. Brain Res 1504: 16-24

88 Kahrstrom C. T. (2013) Epigenetics: Legionella makes its mark on histones. Nat Rev Genet 14: 370

89 Warnault V., Darcq E., Levine A., Barak S. and Ron D. (2013) Chromatin remodeling--a novel strategy to control excessive alcohol drinking. Transl Psychiatry 3: e231

90 Marwick J. A., Kirkham P. A., Stevenson C. S., Danahay H., Giddings J., et al. (2004) Cigarette smoke alters chromatin remodeling and induces proinflammatory genes in rat lungs. Am J Respir Cell Mol Biol 31: 633-642

Rocha P. P., Chaumeil J. and Skok J. A. (2013) Molecular biology. Finding the right partner in a 3D genome. Science 342: 1333-1334

Ronan J. L., Wu W. and Crabtree G. R. (2013) From neural development to cognition: unexpected roles for chromatin. Nat Rev Genet 14: 347-359

Telese F., Gamliel A., Skowronska-Krawczyk D., Garcia-Bassets I. and Rosenfeld M. G. (2013) “Seq-ing” insights into the epigenetics of neuronal gene regulation. Neuron 77: 606-623

Yadon A. N. and Tsukiyama T. (2013) DNA looping-dependent targeting of a chromatin remodeling factor. Cell Cycle 12: 1809-1810

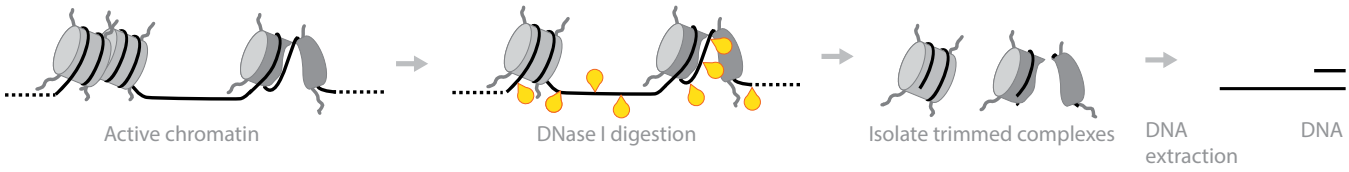
de Wit E. and de Laat W. (2012) A decade of 3C technologies: insights into nuclear organization. Genes Dev 26: 11-24

Sajan S. A. and Hawkins R. D. (2012) Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 13: 59-82

Zentner G. E. and Henikoff S. (2012) Surveying the epigenomic landscape, one base at a time. Genome Biol 13: 250

DNASE L HYPERSENSITIVE SITES SEQUENCING (DNASE-SEQ)

DNase I hypersensitive sites sequencing (DNase-Seq) is based on a well-established DNase I footprinting protocol[91] that was optimized for sequencing[92]. In this method, DNA-protein complexes are treated with DNase I, and the DNA is then extracted and sequenced. Sequences bound by regulatory proteins are protected from DNase I digestion. Deep sequencing provides accurate representation of the location of regulatory proteins in genome. In a variation on this approach, the DNA-protein complexes are stabilized by formaldehyde crosslinking before DNase I digestion. The crosslinking is reversed before DNA purification. In an alternative modification, called GeF-Seq, both the crosslinking and the DNase I digestion are carried out in vivo, within permeabilized cells[93].



Pros	Cons
• *Can detect "open" chromatin*[94] • *No prior knowledge of the sequence or binding protein is required* • *Compared to FAIRE-Seq, has greater sensitivity at promoters*[95]	• *DNase I is sequence-specific and hypersensitive sites might not account for the entire genome* • *Integration of DNase I with ChIP data is necessary to identify and differentiate similar protein-binding sites*

References

Chumsakul O., Nakamura K., Kurata T., Sakamoto T., Hobman J. L., et al. (2013) High-resolution mapping of in vivo genomic transcription factor binding sites using in situ DNase I footprinting and ChIP-seq. DNA Res 20: 325-338

This study describes an improvement and combination of DNase-Seq with ChIP-Seq, called genome footprinting by high throughput sequencing (GeF-Seq).The authors claim GeF-seq provides better alignment due to shorter reads, resulting in higher resolution of DNA-binding factor recognition sites.

Illumina Technology: Genome Analyzer$_{IIx}$

91 Galas D. J. and Schmitz A. (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res 5: 3157-3170

92 Anderson S. (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. Nucleic Acids Res 9: 3015-3027

93 Chumsakul O., Nakamura K., Kurata T., Sakamoto T., Hobman J. L., et al. (2013) High-resolution mapping of in vivo genomic transcription factor binding sites using in situ DNase I footprinting and ChIP-seq. DNA Res 20: 325-338

94 Zentner G. E. and Henikoff S. (2012) Surveying the epigenomic landscape, one base at a time. Genome Biol 13: 250

95 Kumar V., Muratani M., Rayan N. A., Kraus P., Lufkin T., et al. (2013) Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol 31: 615-622

Deng T., Zhu Z. I., Zhang S., Leng F., Cherukuri S., et al. (2013) HMGN1 modulates nucleosome occupancy and DNase I hypersensitivity at the CpG island promoters of embryonic stem cells. Mol Cell Biol 33: 3377-3389

The authors use mouse ESCs and NPCs to study the interplay between histone H1 variants and high-mobility group (HMG) proteins in chromatin remodeling. They use ChIP-Seq and DNase-Seq to elucidate the role of HMGN1 (a HMG protein) in affecting chromatin structure at transcription start sites of promoters.

Illumina Technology: Genome Analyzer$_{IIx}$

Iwata M., Sandstrom R. S., Delrow J. J., Stamatoyannopoulos J. A. and Torok-Storb B. (2013) Functionally and Phenotypically Distinct Subpopulations of Marrow Stromal Cells Are Fibroblast in Origin and Induce Different Fates in Peripheral Blood Monocytes. Stem Cells Dev

Individual cell growth and differentiation is under constant influence by the surrounding tissue and nearby cell types. This study examined marrow stromal cells (MSCs) and their gene expression profiles in comparison to monocyte-derived macrophages that often exist in close proximity to MSCs. Using Illumina sequencing for DNase 1 hypersensitivity mapping, the authors showed a lineage association between two types of MSCs (CD146+,CD146–) and marrow fibroblasts. Subpopulations of CD146+ MSCs were found to increase the expression of genes relevant to hematopoietic regulation upon contact with monocytes, indicating an interaction of fibroblast-macrophage expression.

Illumina Technology: Genome Analyzer$_{IIx}$

Ballare C., Castellano G., Gaveglia L., Althammer S., Gonzalez-Vallinas J., et al. (2013) Nucleosome-driven transcription factor binding and gene regulation. Mol Cell 49: 67-79

Bertucci P. Y., Nacht A. S., Allo M., Rocha-Viegas L., Ballare C., et al. (2013) Progesterone receptor induces bcl-x expression through intragenic binding sites favoring RNA polymerase II elongation. Nucleic Acids Res 41: 6072-6086

Lazarovici A., Zhou T., Shafer A., Dantas Machado A. C., Riley T. R., et al. (2013) Probing DNA shape and methylation state on a genomic scale with DNase I. Proc Natl Acad Sci U S A 110: 6376-6381

Lo K. A., Labadorf A., Kennedy N. J., Han M. S., Yap Y. S., et al. (2013) Analysis of in vitro insulin-resistance models and their physiological relevance to in vivo diet-induced adipose insulin resistance. Cell Rep 5: 259-270

Degner J. F., Pai A. A., Pique-Regi R., Veyrieras J. B., Gaffney D. J., et al. (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. Nature 482: 390-394

Dunowska M., Biggs P. J., Zheng T. and Perrott M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418-424

He H. H., Meyer C. A., Chen M. W., Jordan V. C., Brown M., et al. (2012) Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. Genome Res 22: 1015-1025

Lassen K. S., Schultz H., Heegaard N. H. and He M. (2012) A novel DNAseq program for enhanced analysis of Illumina GAII data: a case study on antibody complementarity-determining regions. N Biotechnol 29: 271-278

Liu M., Li C. L., Stamatoyannopoulos G., Dorschner M. O., Humbert R., et al. (2012) Gammaretroviral Vector Integration Occurs Overwhelmingly Within and Near DNase Hypersensitive Sites. Hum Gene Ther 23: 231-237

Maurano M. T., Humbert R., Rynes E., Thurman R. E., Haugen E., et al. (2012) Systematic localization of common disease-associated variation in regulatory DNA. Science 337: 1190-1195

Neph S., Vierstra J., Stergachis A. B., Reynolds A. P., Haugen E., et al. (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 489: 83-90

Rosseel T., Scheuch M., Hoper D., De Regge N., Caij A. B., et al. (2012) DNase SISPA-next generation sequencing confirms Schmallenberg virus in Belgian field samples and identifies genetic variation in Europe. PLoS One 7: e41967

Wang Y. M., Zhou P., Wang L. Y., Li Z. H., Zhang Y. N., et al. (2012) Correlation between DNase I hypersensitive site distribution and gene expression in HeLa S3 cells. PLoS One 7: e42414

Zhang W., Wu Y., Schnable J. C., Zeng Z., Freeling M., et al. (2012) High-resolution mapping of open chromatin in the rice genome. Genome Res 22: 151-162

Zhang W., Zhang T., Wu Y. and Jiang J. (2012) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. Plant Cell 24: 2719-2731

Boyle A. P., Song L., Lee B. K., London D., Keefe D., et al. (2011) High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. Genome Res 21: 456-464

Stadler M. B., Murr R., Burger L., Ivanek R., Lienert F., et al. (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480: 490-495

McDaniell R., Lee B. K., Song L., Liu Z., Boyle A. P., et al. (2010) Heritable individual-specific and allele-specific chromatin signatures in humans. Science 328: 235-239

Turnbaugh P. J., Quince C., Faith J. J., McHardy A. C., Yatsunenko T., et al. (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. Proc Natl Acad Sci U S A 107: 7503-7508

Audit B., Zaghloul L., Vaillant C., Chevereau G., d'Aubenton-Carafa Y., et al. (2009) Open chromatin encoded in DNA sequence is the signature of 'master' replication origins in human cells. Nucleic Acids Res 37: 6064-6075

Turnbaugh P. J., Ridaura V. K., Faith J. J., Rey F. E., Knight R., et al. (2009) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1: 6ra14

Associated Kits

TruSeq ChIP-Seq kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Preparation Kit

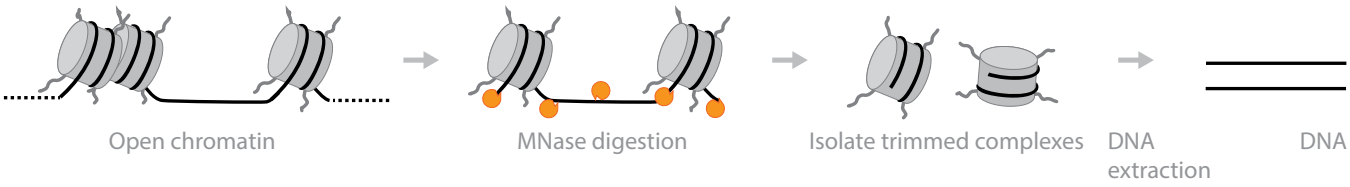
TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

MNASE-ASSISTED ISOLATION OF NUCLEOSOMES SEQUENCING (MAINE-SEQ)

Micrococcal nuclease (MNase)-assisted isolation of nucleosomes sequencing (MAINE-Seq)[96, 97], is a variation on the well-established use of MNase digestion to map nucleosome positions (MNase-Seq)[98]. It is estimated that almost half the genome contains regularly spaced arrays of nucleosomes, which are enriched in active chromatin domains[99]. In MAINE-Seq, genomic DNA is treated with MNase. The DNA from the DNA-protein complexes is then extracted and sequenced. Sequences bound by regulatory proteins are protected from MNase digestion. Deep sequencing provides accurate representation of the location of regulatory proteins in the genome[100]. To identify the regulatory proteins, MNase-Seq can be followed by ChIP (NChIP)[101].



Pros	Cons
• *Can map nucleosomes and other DNA-binding proteins*[102] • *Identifies location of various regulatory proteins in the genome* • *Covers broad range of regulatory sites*	• *MNase sites might not account for the entire genome* • *Does not provide much information about the kind of regulatory elements* • *Integration of MNase with ChIP data is necessary to identify and differentiate similar protein-binding sites*

References

Ballare C., Castellano G., Gaveglia L., Althammer S., Gonzalez-Vallinas J., et al. (2013) Nucleosome-driven transcription factor binding and gene regulation. Mol Cell 49: 67-79

This study combines DNase, ChIP, and MAINE sequencing to understand the effects of chromatin remodeling at hormone-responsive regions and thereby the access of hormone receptors to hormone-responsive elements. The authors report nucleosomal involvement in progesterone receptor binding and hormonal gene regulation.

Illumina Technology: Genome Analyzer$_{IIx}$

96 Cusick M. E., Herman T. M., DePamphilis M. L. and Wassarman P. M. (1981) Structure of chromatin at deoxyribonucleic acid replication forks: prenucleosomal deoxyribonucleic acid is rapidly excised from replicating simian virus 40 chromosomes by micrococcal nuclease. Biochemistry 20: 6648-6658

97 Ponts N., Harris E. Y., Prudhomme J., Wick I., Eckhardt-Ludka C., et al. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res 20: 228-238

98 Schlesinger F., Smith A. D., Gingeras T. R., Hannon G. J. and Hodges E. (2013) De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements. Genome Res 23: 1601-1614

99 Gaffney D. J., McVicker G., Pai A. A., Fondufe-Mittendorf Y. N., Lewellen N., et al. (2012) Controls of nucleosome positioning in the human genome. PLoS Genet 8: e1003036

100 Schones D. E., Cui K., Cuddapah S., Roh T. Y., Barski A., et al. (2008) Dynamic regulation of nucleosome positioning in the human genome. Cell 132: 887-898

101 Boyd-Kirkup J. D., Green C. D., Wu G., Wang D. and Han J. D. (2013) Epigenomics and the regulation of aging. Epigenomics 5: 205-227

102 Zentner G. E. and Henikoff S. (2012) Surveying the epigenomic landscape, one base at a time. Genome Biol 13: 250

Deng T., Zhu Z. I., Zhang S., Leng F., Cherukuri S., et al. (2013) HMGN1 modulates nucleosome occupancy and DNase I hypersensitivity at the CpG island promoters of embryonic stem cells. Mol Cell Biol 33: 3377-3389

Chromatin structure and the interaction of DNA with epigenetic factors and chromatin-remodeling complexes play key roles in regulating gene expression and embryonic differentiation. In this study, the authors applied ChIP-Seq, DNAse I-Seq, and MNase-Seq on an Illumina Genome Analyzer to analyze the organization of nucleosomes in relation to DNase I hypersensitivity and transcription in mouse ESCs. They found that loss of HMG protein HMGN1 affects two important aspects of chromatin organization: altering the nucleosome positioning at the TSS and reducing the number of DNase I hypersensitivity sites.

Illumina Technology: Genome Analyzer$_{IIx}$

Chai X., Nagarajan S., Kim K., Lee K. and Choi J. K. (2013) Regulation of the boundaries of accessible chromatin. PLoS Genet 9: e1003778

Grøntved L., John S., Baek S., Liu Y., Buckley J. R., et al. (2013) C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. EMBO J 32: 1568-1583

Maruyama H., Harwood J. C., Moore K. M., Paszkiewicz K., Durley S. C., et al. (2013) An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis. EMBO Rep 14: 711-717

Nagarajavel V., Iben J. R., Howard B. H., Maraia R. J. and Clark D. J. (2013) Global 'bootprinting' reveals the elastic architecture of the yeast TFIIIB-TFIIIC transcription complex in vivo. Nucleic Acids Res 41: 8135-8143

Nishida H., Katayama T., Suzuki Y., Kondo S. and Horiuchi H. (2013) Base composition and nucleosome density in exonic and intronic regions in genes of the filamentous ascomycetes Aspergillus nidulans and Aspergillus oryzae. Gene 525: 10-May

Tolstorukov M. Y., Sansam C. G., Lu P., Koellhoffer E. C., Helming K. C., et al. (2013) Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. Proc Natl Acad Sci U S A 110: 10165-10170

Allan J., Fraser R. M., Owen-Hughes T. and Keszenman-Pereyra D. (2012) Micrococcal nuclease does not substantially bias nucleosome mapping. J Mol Biol 417: 152-164

Gaffney D. J., McVicker G., Pai A. A., Fondufe-Mittendorf Y. N., Lewellen N., et al. (2012) Controls of nucleosome positioning in the human genome. PLoS Genet 8: e1003036

Guertin M. J., Martins A. L., Siepel A. and Lis J. T. (2012) Accurate prediction of inducible transcription factor binding intensities in vivo. PLoS Genet 8: e1002610

Zhang X., Robertson G., Woo S., Hoffman B. G. and Gottardo R. (2012) Probabilistic Inference for Nucleosome Positioning with MNase-Based or Sonicated Short-Read Data. PLoS ONE 7: e32095

Kent N. A., Adams S., Moorhouse A. and Paszkiewicz K. (2011) Chromatin particle spectrum analysis: a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. Nucleic Acids Res 39: e26

Xi Y., Yao J., Chen R., Li W. and He X. (2011) Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res 21: 718-724

Ponts N., Harris E. Y., Prudhomme J., Wick I., Eckhardt-Ludka C., et al. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res 20: 228-238

Associated Kits

TruSeq ChIP-Seq® Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Preparation Kit

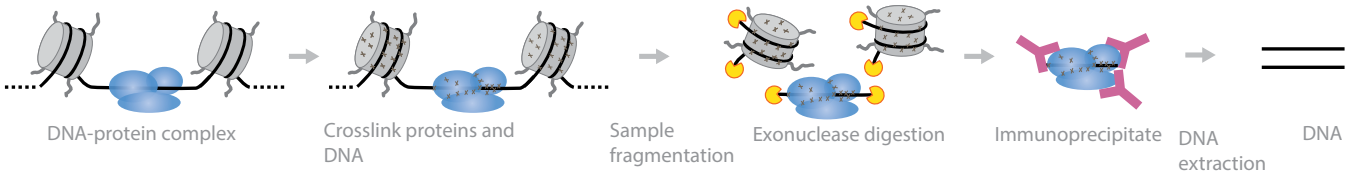
TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

CHROMATIN IMMUNOPRECIPITATION SEQUENCING (CHIP-SEQ)

Chromatin immunoprecipitation sequencing (ChIP-Seq) is a well-established method to map specific protein-binding sites[103] . In this method, DNA-protein complexes are crosslinked *in vivo*. Samples are then fragmented and treated with an exonuclease to trim unbound oligonucleotides. Protein-specific antibodies are used to immunoprecipitate the DNA-protein complex. The DNA is extracted and sequenced, giving high-resolution sequences of the protein-binding sites.



Pros	Cons
• *Base-pair resolution of protein-binding site* • *Specific regulatory factors or proteins can be mapped* • *The use of exonuclease eliminates contamination by unbound DNA*[104]	• *Nonspecific antibodies can dilute the pool of DNA-protein complexes of interest* • *The target protein must be known and able to raise an antibody*

References

Berkseth M., Ikegami K., Arur S., Lieb J. D. and Zarkower D. (2013) TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination. Proc Natl Acad Sci U S A 110: 16033-16038

In an effort to identify targets of the nematode global sexual regulator Transformer 1 (TRA-1), this study applied Illumina sequencing for genome-wide ChIP-Seq analysis of TRA-1 binding sites. The authors identified DNA-binding sites driving male-specific expression patterns and TRA-1 binding sites adjacent to a number of regulatory genes, some of which drive male-specific expression. Overall, the results suggest that TRA-1 mediates sex-specific expression.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Bowman S. K., Simon M. D., Deaton A. M., Tolstorukov M., Borowsky M. L., et al. (2013) Multiplexed Illumina sequencing libraries from picogram quantities of DNA. BMC Genomics 14: 466

This study reports a simple and fast library construction method from sub-nanogram quantities of DNA. This protocol yields conventional libraries with barcodes suitable for multiplexed sample analysis on the Illumina platform. The authors demonstrate the method by constructing a ChIP-Seq library from 100 pg of ChIP DNA that shows equivalent coverage of target regions to a library produced from a larger-scale experiment.

Illumina Technology: HiSeq 2000

103 Solomon M. J., Larsen P. L. and Varshavsky A. (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell 53: 937-947

104 Zentner G. E. and Henikoff S. (2012) Surveying the epigenomic landscape, one base at a time. Genome Biol 13: 250

Kumar V., Muratani M., Rayan N. A., Kraus P., Lufkin T., et al. (2013) Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol 31: 615-622

ChIP-Seq experiments are used to determine the occupation of chromatin by DNA-binding proteins. Data analysis requires detection of binding signals above the background noise, and a common secondary analysis is the prediction of an effect, e.g., expression, from the level of the ChIP-Seq signal. This paper presents algorithms adapted from signal processing theory to solve the two general problems of signal detection and signal estimation from ChIP-Seq data. Using existing data and a new ChIP-Seq data set from an Illumina Genome Analyzer, the two tools DFilter and EFilter are shown to outperform the most commonly used methods in the field, including MACS and Quest.

Illumina Technology: Genome Analyzer$_{IIx}$

Lesch B. J., Dokshin G. A., Young R. A., McCarrey J. R. and Page D. C. (2013) A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. Proc Natl Acad Sci U S A 110: 16061-16066

At conception the zygote is totipotent: incorporating the potential to differentiate into any specialized cell in the body. This study used gene expression profiling and epigenetic regulatory marks (H3K4me3 and H3K37me3) to examine how germ cells change as they progress from differentiated cell to totipotent zygote. The authors used ChIP-Seq and RNA-Seq on the Illumina HiSeq platform for both male and female germ cells at three time points surrounding sex differentiation, meiosis, and post-meiosis. Their results indicate central regulatory genes are maintained in an epigenetically poised state, permitting establishment of totipotency following fertilization.

Illumina Technology: HiSeq2000

Schauer T., Schwalie P. C., Handley A., Margulies C. E., Flicek P., et al. (2013) CAST-ChIP maps cell-type-specific chromatin states in the Drosophila central nervous system. Cell Rep 5: 271-282

Accurate assays for epigenetic markers have been limited by the amount of input material required. This study presents a new assay (CAST-ChIP), based on Illumina sequencing, that allows for characterization of chromatin-associated proteins from specific cell types in complex tissues. The study validates the assay by profiling PolII and H2A.Z across both glia and neurons in Drosophila brain tissue.

Illumina Technology: Genome Analyzer$_{IIx}$

Koldamova R., Schug J., Lefterova M., Cronican A. A., Fitz N. F., et al. (2014) Genome-wide approaches reveal EGR1-controlled regulatory networks associated with neurodegeneration. Neurobiol Dis 63: 107-114

McMullen P. D., Bhattacharya S., Woods C. G., Sun B., Yarborough K., et al. (2014) A map of the PPARalpha transcription regulatory network for primary human hepatocytes. Chem Biol Interact 209: 14-24

Waszak S. M., Kilpinen H., Gschwind A. R., Orioli A., Raghav S. K., et al. (2014) Identification and removal of low-complexity sites in allele-specific analysis of ChIP-seq data. Bioinformatics 30: 165-171

Wolchinsky Z., Shivtiel S., Kouwenhoven E. N., Putin D., Sprecher E., et al. (2014) Angiomodulin is required for cardiogenesis of embryonic stem cells and is maintained by a feedback loop network of p63 and Activin-A. Stem Cell Res 12: 49-59

Biancolella M., B K. F., Tring S., Plummer S. J., Mendoza-Fandino G. A., et al. (2013) Identification and characterization of functional risk variants for colorectal cancer mapping to chromosome 11q23.1. Hum Mol Genet

Chang C. Y., Pasolli H. A., Giannopoulou E. G., Guasch G., Gronostajski R. M., et al. (2013) NFIB is a governor of epithelial-melanocyte stem cell behaviour in a shared niche. Nature 495: 98-102

Chauhan C., Zraly C. B. and Dingwall A. K. (2013) The Drosophila COMPASS-like Cmi-Trr coactivator complex regulates dpp/BMP signaling in pattern formation. Dev Biol 380: 185-198

Jain A., Bacolla A., Del Mundo I. M., Zhao J., Wang G., et al. (2013) DHX9 helicase is involved in preventing genomic instability induced by alternatively structured DNA in human cells. Nucleic Acids Res 41: 10345-10357

Lai C. F., Flach K. D., Alexi X., Fox S. P., Ottaviani S., et al. (2013) Co-regulated gene expression by oestrogen receptor alpha and liver receptor homolog-1 is a feature of the oestrogen response in breast cancer cells. Nucleic Acids Res 41: 10228-10240

Lo K. A., Labadorf A., Kennedy N. J., Han M. S., Yap Y. S., et al. (2013) Analysis of in vitro insulin-resistance models and their physiological relevance to in vivo diet-induced adipose insulin resistance. Cell Rep 5: 259-270

Maehara K., Odawara J., Harada A., Yoshimi T., Nagao K., et al. (2013) A co-localization model of paired ChIP-seq data using a large ENCODE data set enables comparison of multiple samples. Nucleic Acids Res 41: 54-62

Mazzoni E. O., Mahony S., Closser M., Morrison C. A., Nedelec S., et al. (2013) Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. Nat Neurosci 16: 1219-1227

Mendoza-Parra M. A., Van Gool W., Mohamed Saleem M. A., Ceschin D. G. and Gronemeyer H. (2013) A quality control system for profiles obtained by ChIP sequencing. Nucleic Acids Res 41: e196

Neyret-Kahn H., Benhamed M., Ye T., Le Gras S., Cossec J. C., et al. (2013) Sumoylation at chromatin governs coordinated repression of a transcriptional program essential for cell growth and proliferation. Genome Res 23: 1563-1579

Olovnikov I., Ryazansky S., Shpiz S., Lavrov S., Abramov Y., et al. (2013) De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. Nucleic Acids Res 41: 5757-5768

Prickett A. R., Barkas N., McCole R. B., Hughes S., Amante S. M., et al. (2013) Genome-wide and parental allele-specific analysis of CTCF and cohesin DNA binding in mouse brain reveals a tissue-specific binding pattern and an association with imprinted differentially methylated regions. Genome Res 23: 1624-1635

Schmolka N., Serre K., Grosso A. R., Rei M., Pennington D. J., et al. (2013) Epigenetic and transcriptional signatures of stable versus plastic differentiation of proinflammatory gammadelta T cell subsets. Nat Immunol 14: 1093-1100

Sakabe N. J., Aneas I., Shen T., Shokri L., Park S. Y., et al. (2012) Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function. Hum Mol Genet 21: 2194-2204

Tanaka Y., Joshi A., Wilson N. K., Kinston S., Nishikawa S., et al. (2012) The transcriptional programme controlled by Runx1 during early embryonic blood development. Dev Biol 366: 404-419

Tang B., Becanovic K., Desplats P. A., Spencer B., Hill A. M., et al. (2012) Forkhead box protein p1 is a transcriptional repressor of immune signaling in the CNS: implications for transcriptional dysregulation in Huntington disease. Hum Mol Genet 21: 3097-3111

Zhan Q., Fang Y., He Y., Liu H. X., Fang J., et al. (2012) Function annotation of hepatic retinoid x receptor alpha based on genome-wide DNA binding and transcriptome profiling. PLoS One 7: e50013

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Preparation Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

FORMALDEHYDE-ASSISTED ISOLATION OF REGULATORY ELEMENTS (FAIRE-SEQ)

Formaldehyde-assisted isolation of regulatory elements (FAIRE-Seq)[105,106] is based on differences in crosslinking efficiencies between DNA and nucleosomes or sequence-specific DNA-binding proteins. In this method, DNA-protein complexes are briefly crosslinked *in vivo* using formaldehyde. The sample is then lysed and sonicated. After phenol/chloroform extraction, the DNA in the aqueous phase is purified and sequenced. Sequencing provides information for regions of DNA that are not occupied by histones.

Pros	Cons
• *Simple and highly reproducible protocol* • *Does not require antibodies* • *Does not require enzymes, such as DNase or MNase, avoiding the optimization and extra steps necessary for enzymatic processing* • *Does not require a single-cell suspension or nuclear isolation, so it is easily adapted for use on tissue samples*[107]	• *Cannot identify regulatory proteins bound to DNA* • *DNase-Seq may be better at identifying nucleosome-depleted promoters of highly expressed genes*[108]

References

Hilton I. B., Simon J. M., Lieb J. D., Davis I. J., Damania B., et al. (2013) The open chromatin landscape of Kaposi's sarcoma-associated herpesvirus. J Virol 87: 11831-11842

Kaposi's sarcoma-associated herpesvirus (KSHV) is a gammaherpesvirus that, upon infection, remains in a latent state. Histone modifications occupy inactive regions of the latent viral genome. The authors use FAIRE-Seq on the Illumina HiSeq 2000 system to study open chromatin regions in the KSHV genome, allowing them to identify regions of open chromatin in the latent virus. By integrating data on histone modifications, they were able to generate a genome-wide KSHV landscape, which indicated localization of active histone modifications near nucleosome-depleted sites.

Illumina Technology: TruSeq Sample Prep Kit, HiSeq 2000

105 Giresi P. G. and Lieb J. D. (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). Methods 48: 233-239

106 Hogan G. J., Lee C. K. and Lieb J. D. (2006) Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. PLoS Genet 2: e158

107 Simon J. M., Giresi P. G., Davis I. J. and Lieb J. D. (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc 7: 256-267

108 Song L., Zhang Z., Grasfeder L. L., Boyle A. P., Giresi P. G., et al. (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res 21: 1757-1767

Meredith D. M., Borromeo M. D., Deering T. G., Casey B. H., Savage T. K., et al. (2013) Program specificity for Ptf1a in pancreas versus neural tube development correlates with distinct collaborating cofactors and chromatin accessibility. Mol Cell Biol 33: 3166-3179

Transcription factors (TFs) are the drivers of cell development and differentiation. The combined regulatory effects of different TFs allow any factor to play key roles in the different pathways of cell differentiation. This study examined how pancreas-specific transcription factor 1a (Ptf1a) is a critical driver for development of both the pancreas and nervous system. Using Illumina sequencing to perform ChIP-Seq for Ptf1a, FAIRE-Seq to detect open chromatin, and RNA-Seq for expression profiling, the authors characterized Fox and Sox factors as potential lineage-specific modifiers of Ptf1a binding.

Illumina Technology: HiSeq 2000, Genome Analyzer$_{IIx}$

Paul D. S., Albers C. A., Rendon A., Voss K., Stephens J., et al. (2013) Maps of open chromatin highlight cell type-restricted patterns of regulatory sequence variation at hematological trait loci. Genome Res 23: 1130-1141

Genome-wide association studies (GWAS) have discovered many non–protein-coding loci associated with complex traits. However, due to the low resolution of GWAS, the exact location of the causative variant is often not known. In this study, the authors combined GWAS results with FAIRE-Seq to link complex hematopoietic traits to specific functional loci. They found that the majority of candidate functional variants coincided with binding sites of five transcription factors key to regulating megakaryopoiesis, and further found that 76.9% of the candidate regulatory variants affected protein binding at these sites. In conclusion, the combination of GWAS data with high-resolution epigenetic profiling by sequencing is a powerful assay for mapping complex genetic variants.

Illumina Technology: HiSeq 2000, Genome Analyzer$_{IIx}$, Human Gene Expression—BeadArray

Chai X., Nagarajan S., Kim K., Lee K. and Choi J. K. (2013) Regulation of the boundaries of accessible chromatin. PLoS Genet 9: e1003778

Calabrese J. M., Sun W., Song L., Mugford J. W., Williams L., et al. (2012) Site-specific silencing of regulatory elements as a mechanism of X inactivation. Cell 151: 951-963

Simon J. M., Giresi P. G., Davis I. J. and Lieb J. D. (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc 7: 256-267

Paul D. S., Nisbet J. P., Yang T. P., Meacham S., Rendon A., et al. (2011) Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. PLoS Genet 7: e1002139

Ponts N., Harris E. Y., Prudhomme J., Wick I., Eckhardt-Ludka C., et al. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res 20: 228-238

Auerbach R. K., Euskirchen G., Rozowsky J., Lamarre-Vincent N., Moqtaderi Z., et al. (2009) Mapping accessible chromatin regions using Sono-Seq. Proc Natl Acad Sci U S A 106: 14926-14931

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

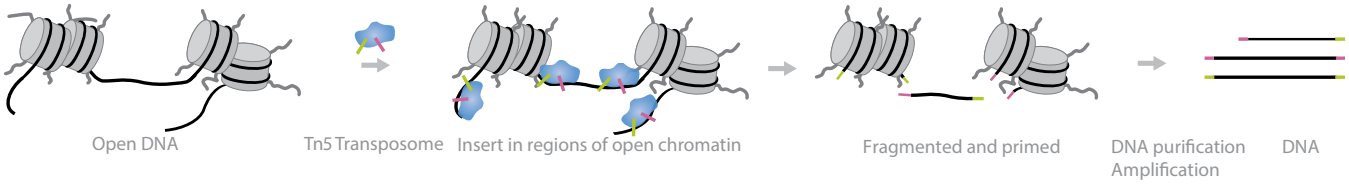
TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

ASSAY FOR TRANSPOSASE-ACCESSIBLE CHROMATIN SEQUENCING (ATAC-SEQ)

Assay for transposase-accessible chromatin using sequencing (ATAC-Seq) is a protocol that utilizes the Epicentre Tn5 transposome[109]. In this method, DNA is incubated with Tn5 transposome, which performs adaptor ligation and fragmentation of open chromatin regions. Deep sequencing of the purified regions provides base-pair resolution of nucleosome-free regions in the genome.



Pros	Cons
• *Two-step protocol with no adaptor ligation steps, gel purification, or crosslink reversal* • *Very high signal to noise ratio compared to FAIRE-Seq*	• *During mechanical sample processing, bound chromatin regions might open and be tagged by the transposome*

References

Buenrostro J. D., Giresi P. G., Zaba L. C., Chang H. Y. and Greenleaf W. J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10: 1213-1218

This is the first paper to describe ATAC-seq as a protocol to study regions of open chromatin. The authors identify the location of DNA-binding proteins in a B-cell line. They demonstrate that the protocol can analyze an individual's T-cell epigenome on a timescale compatible with clinical decision-making.

Illumina Technology: MiSeq, HiSeq 2000

ATAC-Seq enables real-time personal epigenomics.

Associated Kits

EpiGnome™ Methyl-Seq Kit

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

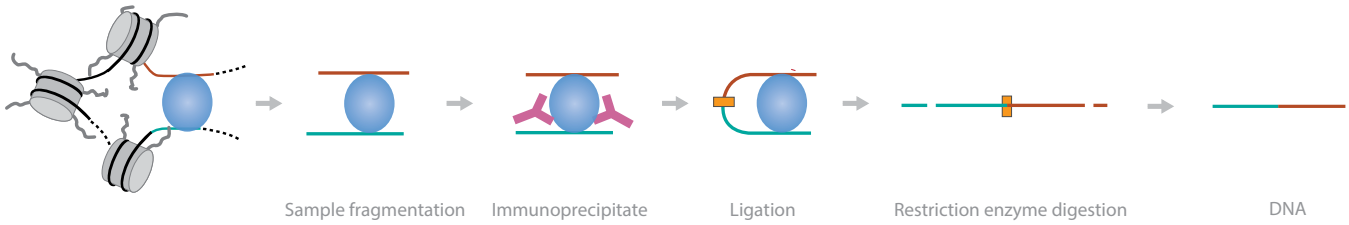
TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

109 Buenrostro J. D., Giresi P. G., Zaba L. C., Chang H. Y. and Greenleaf W. J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10: 1213-1218

CHROMATIN INTERACTION ANALYSIS BY PAIRED-END TAG SEQUENCING (CHIA-PET)

Chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) is a variation of Hi-C that features an immunoprecipitation step to map long-range DNA interactions[110, 111]. In this method, DNA-protein complexes are crosslinked and fragmented. Specific antibodies are used to immunoprecipitate proteins of interest. Specific linkers are ligated to the DNA fragments, which ligate when in proximity. Linkers are then precipitated and digested with an enzyme and the DNA is sequenced. Deep sequencing provides base-pair resolution of ligated fragments. Hi-C and ChIA-PET currently provide the best balance of resolution and reasonable coverage in the human genome to map long-range interactions[112]



Pros	Cons
• *Suitable for detecting a large number of both long-range and short range chromatin interactions globally*[113] • *Studies the interactions made by specific proteins or protein complexes* • *Provides information about DNA interactions aided by regulatory elements* • *Removes background generated during traditional ChIP assays* • *The immunoprecipitation step reduces data complexity*[113]	• *Nonspecific antibodies can pull down unwanted protein complexes and contaminate the pool* • *Linkers can self-ligate, generating ambiguity about true DNA interactions* • *Limited sensitivity; may detect as little as 10% of interactions*[113]

References

DeMare L. E., Leng J., Cotney J., Reilly S. K., Yin J., et al. (2013) The genomic landscape of cohesin-associated chromatin interactions. Genome Res 23: 1224-1234

Knockdown of cohesin in ESCs results in aberrant gene expression and loss of pluripotency. Cohesin works to stabilize DNA by forming loops between distant-acting enhancers and their target promoters. The authors studied cohesin interaction in the developing limb using ChIA-PET, RNA-Seq, and ChIP-Seq analysis performed on a HiSeq 2000 system. They report tissue-specific enhancer-promoter interactions involving cohesin and the insulator protein CTCF. They also identified interactions that are maintained for tissue-specific activation or repression during development.

Illumina Technology: TruSeq Sample Prep Kit, HiSeq 2000

110 Li G., Fullwood M. J., Xu H., Mulawadi F. H., Velkov S., et al. (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol 11: R22

111 Fullwood M. J., Liu M. H., Pan Y. F., Liu J., Xu H., et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462: 58-64

112 Dekker J., Marti-Renom M. A. and Mirny L. A. (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14: 390-403

113 Sajan S. A. and Hawkins R. D. (2012) Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 13: 59-82

Stadhouders R., Kolovos P., Brouwer R., Zuin J., van den Heuvel A., et al. (2013) Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. Nat Protoc 8: 509-524

This paper presents an assay for multiplexed chromosome conformation capture sequencing (3C-Seq) using an Illumina HiSeq 2000 system. This high-throughput assay outperforms PCR-based methods for ease of multiplexing, and outperforms 5C and Hi-C methods in terms of cost and ease of analysis. The preparation of multiplexed 3C-Seq libraries can be performed by any investigator with basic skills in molecular biology techniques, and the data analysis requires only basic expertise in bioinformatics.

Illumina Technology: HiSeq 2000

Li G., Ruan X., Auerbach R. K., Sandhu K. S., Zheng M., et al. (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148: 84-98

Zhang J., Poh H. M., Peh S. Q., Sia Y. Y., Li G., et al. (2012) ChIA-PET analysis of transcriptional chromatin interactions. Methods 58: 289-299

Tan S. K., Lin Z. H., Chang C. W., Varang V., Chng K. R., et al. (2011) AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. EMBO J 30: 2569-2581

Fullwood M. J., Han Y., Wei C. L., Ruan X. and Ruan Y. (2010) Chromatin interaction analysis using paired-end tag sequencing. Curr Protoc Mol Biol Chapter 21: Unit 21 15 21-25

Li G., Fullwood M. J., Xu H., Mulawadi F. H., Velkov S., et al. (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol 11: R22

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

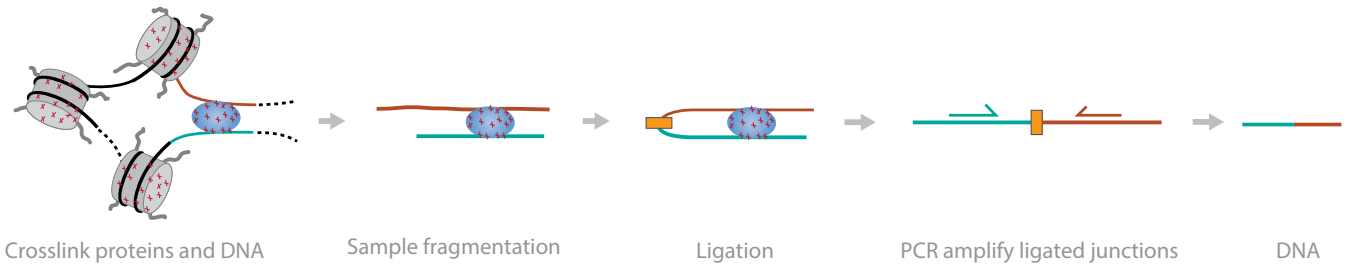
Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate® Pair Kit

CHROMATIN CONFORMATION CAPTURE (HI-C/3C-SEQ)

Chromatin conformation capture sequencing (Hi-C)[114] or 3C-Seq[115] is used to analyze chromatin interactions. In this method, DNA-protein complexes are crosslinked using formaldehyde. The sample is fragmented and DNA ligated and digested. The resulting DNA fragments are PCR-amplified and sequenced. Deep sequencing provides base-pair resolution of ligated fragments.



Pros	Cons
• *Allows detection of long-range DNA interactions* • *High-throughput method*	• *Detection may result from random chromosomal collisions* • *3C PCR is difficult and requires careful controls and experimental design* • *Needs further confirmation of interaction* • *Due to multiple steps, the method requires large amounts of starting material*

References

Burton J. N., Adey A., Patwardhan R. P., Qiu R., Kitzman J. O., et al. (2013) Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol 31: 1119-1125

The authors integrate shotgun fragment and short insert mate-pair sequences with Hi-C data to generate assemblies for human, mouse, and Drosophila genomes. The paper reports a bioinformatics tool used to compute the assemblies: ligating adjacent chromatin enables scaffolding in situ (LACHESIS).

Illumina Technology: HiSeq 2000

Jin F., Li Y., Dixon J. R., Selvaraj S., Ye Z., et al. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503: 290-294

Cis-acting regulatory elements in the genome interact with their target gene promoter by transcription factors bringing the two locations close in the three-dimensional conformation of the chromatin. In this study, the chromosome conformation is studied by a genome-wide analysis method (Hi-C) using the Illumina HiSeq 2000 system. The authors determined over one million long-range chromatin interactions in human fibroblasts. In addition, they characterized the dynamics of promoter-enhancer contacts after TNF-alpha signaling and discovered pre-existing chromatin looping with TNF-alpha–responsive enhancers, suggesting the three-dimensional chromatin conformation may be stable over time.

Illumina Technology: HiSeq 2000

114 Lieberman-Aiden E., van Berkum N. L., Williams L., Imakaev M., Ragoczy T., et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289-293

115 Duan Z., Andronescu M., Schutz K., Lee C., Shendure J., et al. (2012) A genome-wide 3C-method for characterizing the three-dimensional architectures of genomes. Methods 58: 277-288

Nagano T., Lubling Y., Stevens T. J., Schoenfelder S., Yaffe E., et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502: 59-64

Belton J. M., McCord R. P., Gibcus J. H., Naumova N., Zhan Y., et al. (2012) Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 58: 268-276

Demichelis F., Setlur S. R., Banerjee S., Chakravarty D., Chen J. Y., et al. (2012) Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk. Proc Natl Acad Sci U S A 109: 6686-6691

Dixon J. R., Selvaraj S., Yue F., Kim A., Li Y., et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485: 376-380

Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA; (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9: 999-1003

Imakaev M., Fudenberg G., McCord R. P., Naumova N., Goloborodko A., et al. (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9: 999-1003

Irimia M., Tena J. J., Alexis M. S., Fernandez-Minan A., Maeso I., et al. (2012) Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. Genome Res 22: 2356-2367

Lan X., Witt H., Katsumura K., Ye Z., Wang Q., et al. (2012) Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. Nucleic Acids Res 40: 7690-7704

Verma-Gaur J., Torkamani A., Schaffer L., Head S. R., Schork N. J., et al. (2012) Noncoding transcription within the Igh distal V(H) region at PAIR elements affects the 3D structure of the Igh locus in pro-B cells. Proc Natl Acad Sci U S A 109: 17004-17009

Zhang Y., McCord R. P., Ho Y. J., Lajoie B. R., Hildebrand D. G., et al. (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 148: 908-921

Yaffe E. and Tanay A. (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet 43: 1059-1065

Yaffe E, Tanay A; (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet 43: 1059-65

Lieberman-Aiden E., van Berkum N. L., Williams L., Imakaev M., Ragoczy T., et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289-293

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

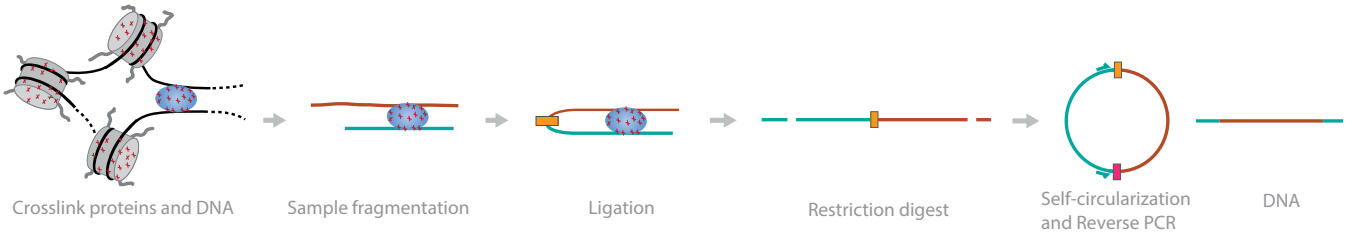
Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

CIRCULAR CHROMATIN CONFORMATION CAPTURE (4-C OR 4C-SEQ)

Circular chromatin conformation capture (4-C)[116], also called 4C-Seq, is a method similar to 3-C and is sometimes called circular 3C. It allows the unbiased detection of all genomic regions that interact with a particular region of interest[117]. In this method, DNA-protein complexes are crosslinked using formaldehyde. The sample is fragmented, and the DNA is ligated and digested. The resulting DNA fragments self-circularize, followed by reverse PCR and sequencing. Deep sequencing provides base-pair resolution of ligated fragments.



Pros	Cons
• *4C is the preferred strategy to assess the DNA contact profile of individual genomic sites.* • *Highly reproducible data*	• *Will miss local interactions (< 50 kb) from the region of interest* • *The large circles do not PCR efficiently*

References

de Wit E., Bouwman B. A., Zhu Y., Klous P., Splinter E., et al. (2013) The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature 501: 227-231

Transcriptional regulation is influenced by the availability of specific transcription factors, but the evidence is increasing for the substantial importance of chromatin conformation within the nucleus. In this study, Illumina sequencing is used to analyze chromatin conformation by a genome-wide assay (4-C) demonstrating, along with ChIP-Seq data, that inactive chromatin is disorganized in PSC nuclei. In contrast to inactive chromatin, promoters are seen to engage in contacts between topological domains in a tissue-dependent manner, while enhancers have a more tissue-restricted interaction. The authors hypothesize that the chromatin interactions enhance the robustness of the pluripotent state.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

116 Zhao Z., Tavoosidana G., Sjolinder M., Gondor A., Mariano P., et al. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38: 1341-1347

117 Sajan S. A. and Hawkins R. D. (2012) Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 13: 59-82

Holwerda S. J., van de Werken H. J., Ribeiro de Almeida C., Bergen I. M., de Bruijn M. J., et al. (2013) Allelic exclusion of the immunoglobulin heavy chain locus is independent of its nuclear localization in mature B cells. Nucleic Acids Res 41: 6905-6916

Chromatin conformation is one of many mechanisms for regulating gene expression. In developing B cells, the immunoglobulin heavy chain (IgH) locus undergoes a scheduled genomic rearrangement of the V, D, and J gene segments. In this study, an allele-specific chromosome conformation capture sequencing technique (4C-Seq) was applied to unambiguously follow the individual IgH alleles in mature B lymphocytes. The authors found that IgH adopts a lymphoid-specific nuclear location, and in mature B cells the distal VH regions of both IgH alleles position themselves away from active chromatin.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq 2000

Wei Z., Gao F., Kim S., Yang H., Lyu J., et al. (2013) Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. Cell Stem Cell 13: 36-47

PSCs are capable of differentiation into diverse cell types. The maintenance of pluripotency and the induction of differentiation are both highly regulated processes. This study examined the epigenetic mechanisms underlying reprogramming of PSCs. Using circular chromosome conformation capture with Illumina HiSeq sequencing technology (4C-Seq), the authors profiled the PSC-specific long-range chromosomal interactions during reprogramming to induced PSCs. The high-resolution genome-wide interaction map and a well-designed experimental setup allowed the authors to show evidence for a functional role of Kruppel-like factor 4 (Klf4) in facilitating long-range interactions.

Illumina Technology: Genome Analyzer$_{IIx}$, HiSeq2000

Chaumeil J., Micsinai M., Ntziachristos P., Roth D. B., Aifantis I., et al. (2013) The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. Nat Commun 4: 2231

Delpretti S., Montavon T., Leleu M., Joye E., Tzika A., et al. (2013) Multiple Enhancers Regulate Hoxd Genes and the Hotdog LncRNA during Cecum Budding. Cell Rep 5: 137-150

Denholtz M., Bonora G., Chronis C., Splinter E., de Laat W., et al. (2013) Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization. Cell Stem Cell 13: 602-616

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Medvedovic J., Ebert A., Tagoh H., Tamir I. M., Schwickert T. A., et al. (2013) Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. Immunity 39: 229-244

Yamane A., Robbiani D. F., Resch W., Bothmer A., Nakahashi H., et al. (2013) RPA Accumulation during Class Switch Recombination Represents 5'-3' DNA-End Resection during the S-G2/M Phase of the Cell Cycle. Cell Rep 3: 138-147

Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K. R., et al. (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484: 69-74

Kovalchuk A. L., Ansarah-Sobrinho C., Hakim O., Resch W., Tolarova H., et al. (2012) Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. Proc Natl Acad Sci U S A 109: 10972-10977

Rocha P. P., Micsinai M., Kim J. R., Hewitt S. L., Souza P. P., et al. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47: 873-885

Sexton T., Kurukuti S., Mitchell J. A., Umlauf D., Nagano T., et al. (2012) Sensitive detection of chromatin coassociations using enhanced chromosome conformation capture on chip. Nat Protoc 7: 1335-1350

Splinter E., de Wit E., van de Werken H. J., Klous P. and de Laat W. (2012) Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods 58: 221-230

van de Werken H. J., Landan G., Holwerda S. J., Hoichman M., Klous P., et al. (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9: 969-972

Hakim O., Sung M. H., Voss T. C., Splinter E., John S., et al. (2011) Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. Genome Res 21: 697-706

Oliveira T., Resch W., Jankovic M., Casellas R., Nussenzweig M. C., et al. (2011) Translocation capture sequencing: A method for high throughput mapping of chromosomal rearrangements. J Immunol Methods 375: 176-181

Robyr D., Friedli M., Gehrig C., Arcangeli M., Marin M., et al. (2011) Chromosome conformation capture uncovers potential genome-wide interactions between human conserved non-coding sequences. PLoS ONE 6: e17634

Rodley C. D., Pai D. A., Mills T. A., Engelke D. R. and O'Sullivan J. M. (2011) tRNA gene identity affects nuclear positioning.

PLoS One 6: e29267

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

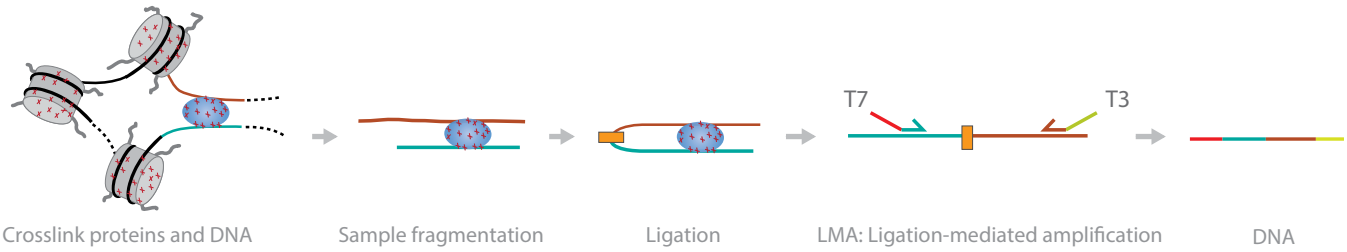
Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

CHROMATIN CONFORMATION CAPTURE CARBON COPY (5-C)

Chromatin conformation capture carbon copy (5-C)[118] allows concurrent determination of interactions between multiple sequences and is a high-throughput version of 3-C[119]. In this method, DNA-protein complexes are crosslinked using formaldehyde. The sample is fragmented and the DNA ligated and digested. The resulting DNA fragments are amplified using ligation-mediated PCR and sequenced. Deep sequencing provides base-pair resolution of ligated fragments.



Pros	Cons
• *Different from 4-C, 5C provides a matrix of interaction frequencies for many pairs of sites* • *Can be used to reconstruct the (average) 3D conformation of larger genomic regions*[120]	• *Detection may not necessarily mean an interaction, resulting from random chromosomal collisions* • *Needs further confirmation of interaction* • *Cannot scale to genome-wide studies that would require large amount of primers*

References

Nora E. P., Lajoie B. R., Schulz E. G., Giorgetti L., Okamoto I., et al. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485: 381-385

The authors use 5-C to analyze regulation of Xist, a non–protein coding transcript that is controlled by X-inactivation center (Xic) to initiate X chromosome inactivation in mouse. They identify a regulatory region of Xist antisense unit that produces a long overriding RNA.

Illumina Technology: Genome Analyzer$_{IIx}$

Sanyal A., Lajoie B. R., Jain G. and Dekker J. (2012) The long-range interaction landscape of gene promoters. Nature 489: 109-113

Associated Kits

TruSeq ChIP-Seq Kit

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

118 Dostie J. and Dekker J. (2007) Mapping networks of physical interactions between genomic elements using 5C technology. Nat Protoc 2: 988-1002

119 Sajan S. A. and Hawkins R. D. (2012) Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 13: 59-82

120 de Wit E. and de Laat W. (2012) A decade of 3C technologies: insights into nuclear organization. Genes Dev 26: 11-24

SEQUENCE REARRANGEMENTS

A growing body of evidence suggests that somatic genomic rearrangements, such as retrotransposition and copy number variants (CNVs), are relatively common in healthy individuals[121,122,123]. Cancer genomes are also known to contain numerous complex rearrangements[124]. While many of these rearrangements can be detected during routine next-generation sequencing, specific techniques are available to study rearrangements such as transposable elements.

Transposable genetic elements (TEs) comprise a vast array of DNA sequences with the ability to move to new sites in genomes either directly by a cut-and-paste mechanism (transposons) or indirectly through an RNA intermediate (retrotransposons)[125]. TEs make up about 66-69% of the human genome[126] and play roles in ageing, cancers, brain, development, embryogenesis, and phenotypic variation in populations and evolution. TEs played a major role in dynamic arrangement of the sex determining region over evolution, giving us distinct X and Y chromosomes[127].

Along with sequence rearrangements by TEs, chromosome and centromere rearrangements can also lead to multiple diseases and disorders[128]. Prenatal diagnostics to study rearrangements predict genetic abnormalities in the fetus. The role of specific TEs and the primary mechanism of chromosome and centromere rearrangements have yet to be elucidated; studying them will help understand their roles.

Transposable elements involved in the evolution of sex chromosomes.

121 O'Huallachain M., Weissman S. M. and Snyder M. P. (2013) The variable somatic genome. Cell Cycle 12: 5-6

122 Macosko E. Z. and McCarroll S. A. (2013) Genetics. Our fallen genomes. Science 342: 564-565

123 McConnell M. J., Lindberg M. R., Brennand K. J., Piper J. C., Voet T., et al. (2013) Mosaic copy number variation in human neurons. Science 342: 632-637

124 Malhotra A., Lindberg M., Faust G. G., Leibowitz M. L., Clark R. A., et al. (2013) Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. Genome Res 23: 762-776

125 Fedoroff N. V. (2012) Presidential address. Transposable elements, epigenetics, and genome evolution. Science 338: 758-767

126 Grandi F. C. and An W. (2013) Non-LTR retrotransposons and microsatellites: Partners in genomic variation. Mob Genet Elements 3: e25674

127 Gschwend A. R., Weingartner L. A., Moore R. C. and Ming R. (2012) The sex-specific region of sex chromosomes in animals and plants. Chromosome Res 20: 57-69

128 Chiarle R. (2013) Translocations in normal B cells and cancers: insights from new technical approaches. Adv Immunol 117: 39-71

Reviews

Bunting S. F. and Nussenzweig A. (2013) End-joining, translocations and cancer. Nat Rev Cancer 13: 443-454

Chiarle R. (2013) Translocations in normal B cells and cancers: insights from new technical approaches. Adv Immunol 117: 39-71

Gifford W. D., Pfaff S. L. and Macfarlan T. S. (2013) Transposable elements as genetic regulatory substrates in early development. Trends Cell Biol 23: 218-226

Grandi F. C. and An W. (2013) Non-LTR retrotransposons and microsatellites: Partners in genomic variation. Mob Genet Elements 3: e25674

van Opijnen T. and Camilli A. (2013) Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol 11: 435-442

Burns K. H. and Boeke J. D. (2012) Human transposon tectonics. Cell 149: 740-752

Fedoroff N. V. (2012) Presidential address. Transposable elements, epigenetics, and genome evolution. Science 338: 758-767

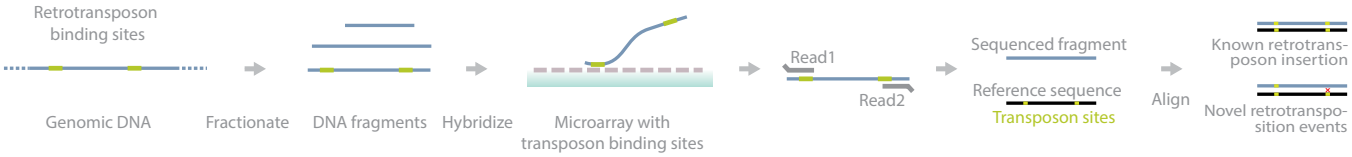
Gschwend A. R., Weingartner L. A., Moore R. C. and Ming R. (2012) The sex-specific region of sex chromosomes in animals and plants. Chromosome Res 20: 57-69

Hancks D. C. and Kazazian H. H., Jr. (2012) Active human retrotransposons: variation and disease. Curr Opin Genet Dev 22: 191-203

Febrer M., McLay K., Caccamo M., Twomey K. B. and Ryan R. P. (2011) Advances in bacterial transcriptome and transposon insertion-site profiling using second-generation sequencing. Trends Biotechnol 29: 586-594

RETROTRANSPOSON CAPTURE SEQUENCING (RC-SEQ)

Retrotransposon capture sequencing (RC-Seq) is a high-throughput protocol to map and study retrotransposon insertions[129]. In this method, after genomic DNA is fractionated, retrotransposon binding sites on DNA hybridize to transposon binding sites on a microarray. Deep sequencing provides accurate information that can be aligned to a reference sequence to discover novel retrotransposition events.



Pros	Cons
• *Ability to clearly identify and detect novel retrotransposition events* • *Can specifically study transposon binding sites of interest* • *High-throughput protocol*	• *Different types of MEI require separate PCR experiments with different primers*[130] • *Hybridization errors can lead to sequencing unwanted DNA fragments* • *PCR biases can underrepresent GC-rich templates* • *Similar transposition binding sites can lead to sequence ambiguity and detection for a transposition event*

References

Shukla R., Upton K. R., Munoz-Lopez M., Gerhardt D. J., Fisher M. E., et al. (2013) Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell 153: 101-111

LINE-1 (L1) retrotransposons are mobile genetic elements comprising ~17% of the human genome. To investigate the significance of novel L1 insertions in cancer, this study used RC-Seq on an Illumina HiSeq 2000 system for 19 hepatocellular carcinoma (HCC) and colorectal cancers (MCC). From these data, the authors identified novel L1 insertion events: each individual genome contained on average 244 non-reference L1 insertions. Forty-five non-reference insertions were annotated as tumor-specific and three of these insertions coincided with strong inhibition of the tumor suppressor MCC. These data provide substantial evidence for L1-mediated retrotransposition playing a role in HCC development.

Illumina Technology: HiSeq 2000

129 Baillie J. K., Barnett M. W., Upton K. R., Gerhardt D. J., Richmond T. A., et al. (2011) Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479: 534-537

130 Xing J., Witherspoon D. J. and Jorde L. B. (2013) Mobile element biology: new possibilities with high-throughput sequencing. Trends Genet 29: 280-289

Solyom S., Ewing A. D., Rahrmann E. P., Doucet T., Nelson H. H., et al. (2012) Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res 22: 2328-2338

Baillie J. K., Barnett M. W., Upton K. R., Gerhardt D. J., Richmond T. A., et al. (2011) Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479: 534-537

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

TRANSPOSON SEQUENCING (TN-SEQ) OR INSERTION SEQUENCING (INSEQ)

Transposon sequencing (Tn-Seq) or insertion sequencing (INSeq) accurately determines quantitative genetic interactions[131]. In this method, a transposon with flanking MmeI digestion sites is transposed into bacteria which, after culturing, can help detect the frequency of mutations within the transposon. After MmeI digestion and subsequent adapter ligation, PCR amplification and sequencing can provide information about the transposon insertion sites.

Pros	Cons
• *Can study mutational frequency of transposons* • *Method can be used to deduce fitness of genes within microorganisms* • *Protocol is robust, reproducible, and sensitive*	• *Limited to bacterial studies* • *Errors during PCR amplification can lead to inaccurate sequence reads*

References

Dong T. G., Ho B. T., Yoder-Himes D. R. and Mekalanos J. J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. Proc Natl Acad Sci U S A 110: 2623-2628.

T6SS is an important protein for bacterial competition; however, T6SS-dependent effector and immunity proteins have not yet been determined. In this study, the authors use Tn-Seq to identify these proteins in *Vibrio cholerae*.

Illumina Technology: HiSeq 2000

Troy E. B., Lin T., Gao L., Lazinski D. W., Camilli A., et al. (2013) Understanding barriers to Borrelia burgdorferi dissemination during infection using massively parallel sequencing. Infect Immun 81: 2347-2357

Infection by *Borrelia burgdorferi* can cause chronic infections of skin, heart, joints, and the central nervous system of infected mammalian hosts. In this study, the authors characterized the population dynamics of mixed populations of *B. burgdorferi* during infection in a mouse model. Using Tn-Seq based on Illumina technology, they mapped the compositions of *B. burgdorferi* at both the injection site and in distal tissues. The authors found that the infection site was a population bottleneck that significantly altered the composition of the population; however, no such bottleneck was observed in colonization of distal tissues.

Illumina Technology: Genome Analyzer$_{IIx}$

131 van Opijnen T., Bodi K. L. and Camilli A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 6: 767-772

Murray S. M., Panis G., Fumeaux C., Viollier P. H. and Howard M. (2013) Computational and genetic reduction of a cell cycle to its simplest, primordial components. PLoS Biol 11: e1001749

Sarmiento F., Mrázek J. and Whitman W. B. (2013) Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon Methanococcus maripaludis. Proc Natl Acad Sci U S A 110: 4726-4731

Skurnik D., Roux D., Cattoir V., Danilchanka O., Lu X., et al. (2013) Enhanced in vivo fitness of carbapenem-resistant oprD mutants of Pseudomonas aeruginosa revealed through high-throughput sequencing. Proc Natl Acad Sci U S A 110: 20747-20752

Mann B., van Opijnen T., Wang J., Obert C., Wang Y. D., et al. (2012) Control of virulence by small RNAs in Streptococcus pneumoniae. PLoS Pathog 8: e1002788

Qi X., Daily K., Nguyen K., Wang H., Mayhew D., et al. (2012) Retrotransposon profiling of RNA polymerase III initiation sites. Genome Res 22: 681-692

van Opijnen T. and Camilli A. (2012) A fine scale phenotype-genotype virulence map of a bacterial pathogen. Genome Res 22: 2541-2551

Zomer A., Burghout P., Bootsma H. J., Hermans P. W. and van Hijum S. A. (2012) ESSENTIALS: Software for Rapid Analysis of High Throughput Transposon Insertion Sequencing Data. PLoS ONE 7: e43012

Goodman A. L., Wu M. and Gordon J. I. (2011) Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. Nat Protoc 6: 1969-1980

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

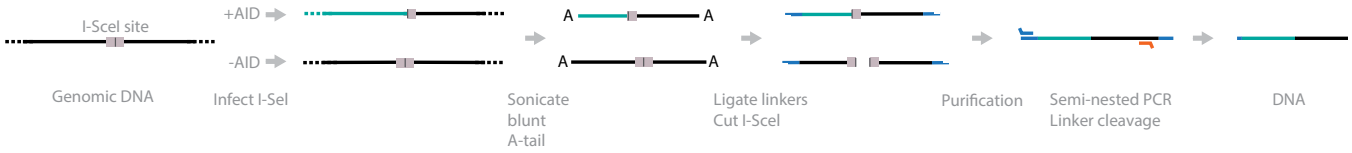
Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

TRANSLOCATION-CAPTURE SEQUENCING (TC-SEQ)

Translocation-capture sequencing (TC-Seq) is a method developed to study chromosomal rearrangements and translocations[132]. In this method, cells are infected with retrovirus expressing I-SceI sites in cells with and without activation-induced cytidine deaminase (AICDA or AID) protein. Genomic DNA from cells is sonicated, linker-ligated, purified, and amplified via semi-nested LM-PCR. The linker is then cleaved and the DNA is sequenced. Any AID-dependent chromosomal rearrangement will be amplified by LM-PCR, while AID-independent translocations will be discarded.



Pros	Cons
• *Can study chromosomal translocations within a given model or environment* • *Random sonication generates unique linker ligation points, and deep sequencing allows reading through rearrangement breakpoints*	• *PCR amplification errors* • *Non-linear PCR amplification can lead to biases affecting reproducibility* • *PCR biases can underrepresent GC-rich templates*

References

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Programmed DNA rearrangement in lymphocytes is initiated by AID protein. The overexpression of AID is associated with cancer, but overexpression of AID alone is insufficient to produce malignancy. This study examines the roles of AID and tumor suppressor p53-binding protein 1 (53BP1) in combination. The results show that the combination of 53BP1 deficiency and AID deregulation increases the rate of rearrangements and results in B cell lymphoma in a mouse model. The rate of rearrangements and CNVs are studied using the Illumina Genome Analyzer.

Illumina Technology: Genome Analyzer$_{IIx}$

132 Klein I. A., Resch W., Jankovic M., Oliveira T., Yamane A., et al. (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147: 95-106

Kovalchuk A. L., Ansarah-Sobrinho C., Hakim O., Resch W., Tolarova H., et al. (2012) Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. Proc Natl Acad Sci U S A 109: 10972-10977

Oliveira T. Y., Resch W., Jankovic M., Casellas R., Nussenzweig M. C., et al. (2012) Translocation capture sequencing: a method for high throughput mapping of chromosomal rearrangements. J Immunol Methods 375: 176-181

Qi X., Daily K., Nguyen K., Wang H., Mayhew D., et al. (2012) Retrotransposon profiling of RNA polymerase III initiation sites. Genome Res 22: 681-692

Rocha P. P., Micsinai M., Kim J. R., Hewitt S. L., Souza P. P., et al. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47: 873-885

Klein I. A., Resch W., Jankovic M., Oliveira T., Yamane A., et al. (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147: 95-106

Associated Kits

TruSeq Nano DNA Sample Prep Kit

TruSeq DNA Sample Prep Kit

TruSeq DNA PCR-Free Sample Prep Kit

Nextera DNA Sample Prep Kit

Nextera XT DNA Sample Prep Kit

Nextera Mate Pair Kit

Nextera Rapid Capture Exome/Custom Enrichment Kit

BIBLIOGRAPHY

Adelman K. and Lis J. T. (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet 13: 720-731

Adey A. and Shendure J. (2012) Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. Genome Res 22: 1139-1143

Allan J., Fraser R. M., Owen-Hughes T. and Keszenman-Pereyra D. (2012) Micrococcal nuclease does not substantially bias nucleosome mapping. J Mol Biol 417: 152-164

Anderson S. (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. Nucleic Acids Res 9: 3015-3027

Audit B., Zaghloul L., Vaillant C., Chevereau G., d'Aubenton-Carafa Y., et al. (2009) Open chromatin encoded in DNA sequence is the signature of 'master' replication origins in human cells. Nucleic Acids Res 37: 6064-6075

Auerbach R. K., Euskirchen G., Rozowsky J., Lamarre-Vincent N., Moqtaderi Z., et al. (2009) Mapping accessible chromatin regions using Sono-Seq. Proc Natl Acad Sci U S A 106: 14926-14931

Baillie J. K., Barnett M. W., Upton K. R., Gerhardt D. J., Richmond T. A., et al. (2011) Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479: 534-537

Ball M. P., Li J. B., Gao Y., Lee J. H., LeProust E. M., et al. (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 27: 361-368

Ballare C., Castellano G., Gaveglia L., Althammer S., Gonzalez-Vallinas J., et al. (2013) Nucleosome-driven transcription factor binding and gene regulation. Mol Cell 49: 67-79

Baslan T., Kendall J., Rodgers L., Cox H., Riggs M., et al. (2012) Genome-wide copy number analysis of single cells. Nat Protoc 7: 1024-1041

Becker A. H., Oh E., Weissman J. S., Kramer G. and Bukau B. (2013) Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. Nat Protoc 8: 2212-2239

Belton J. M., McCord R. P., Gibcus J. H., Naumova N., Zhan Y., et al. (2012) Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 58: 268-276

Berkseth M., Ikegami K., Arur S., Lieb J. D. and Zarkower D. (2013) TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination. Proc Natl Acad Sci U S A 110: 16033-16038

Bertucci P. Y., Nacht A. S., Allo M., Rocha-Viegas L., Ballare C., et al. (2013) Progesterone receptor induces bcl-x expression through intragenic binding sites favoring RNA polymerase II elongation. Nucleic Acids Res 41: 6072-6086

Bhargava V., Ko P., Willems E., Mercola M. and Subramaniam S. (2013) Quantitative transcriptomics using designed primer-based amplification. Sci Rep 3: 1740

Bhargava V., Head S. R., Ordoukhanian P., Mercola M. and Subramaniam S. (2014) Technical variations in low-input RNA-seq methodologies. Sci Rep 4: 3678

Bian C. and Yu X. (2013) PGC7 suppresses TET3 for protecting DNA methylation. Nucleic Acids Res Biancolella M., B K. F., Tring S., Plummer S. J., Mendoza-Fandino G. A., et al. (2013) Identification and characterization of functional risk variants for colorectal cancer mapping to chromosome 11q23.1. Hum Mol Genet

Blainey P. C. (2013) The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 37: 407-427

Blaschke K., Ebata K. T., Karimi M. M., Zepeda-Martinez J. A., Goyal P., et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500: 222-226

Bock C., Tomazou E. M., Brinkman A. B., Muller F., Simmer F., et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114

Bock C., Kiskinis E., Verstappen G., Gu H., Boulting G., et al. (2011) Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 144: 439-452

Booth M. J., Branco M. R., Ficz G., Oxley D., Krueger F., et al. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 336: 934-937

Booth M. J., Ost T. W., Beraldi D., Bell N. M., Branco M. R., et al. (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat Protoc 8: 1841-1851

Böttcher R., Amberg R., Ruzius F. P., Guryev V., Verhaegh W. F., et al. (2012) Using a priori knowledge to align sequencing reads to their exact genomic position. Nucleic Acids Res 40: e125

Bowman S. K., Simon M. D., Deaton A. M., Tolstorukov M., Borowsky M. L., et al. (2013) Multiplexed Illumina sequencing libraries from picogram quantities of DNA. BMC Genomics 14: 466

Boyd-Kirkup J. D., Green C. D., Wu G., Wang D. and Han J. D. (2013) Epigenomics and the regulation of aging. Epigenomics 5: 205-227

Boyle A. P., Song L., Lee B. K., London D., Keefe D., et al. (2011) High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. Genome Res 21: 456-464

Brinkman A. B., Simmer F., Ma K., Kaan A., Zhu J., et al. (2010) Whole-genome DNA methylation profiling using MethylCap-seq. Methods 52: 232-236

Brinkman A. B., Gu H., Bartels S. J., Zhang Y., Matarese F., et al. (2012) Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 22: 1128-1138

Broughton J. P. and Pasquinelli A. E. (2013) Identifying Argonaute binding sites in Caenorhabditis elegans using iCLIP. Methods 63: 119-125

Buenrostro J. D., Giresi P. G., Zaba L. C., Chang H. Y. and Greenleaf W. J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10: 1213-1218

Bunting S. F. and Nussenzweig A. (2013) End-joining, translocations and cancer. Nat Rev Cancer 13: 443-454

Burns K. H. and Boeke J. D. (2012) Human transposon tectonics. Cell 149: 740-752

Burton J. N., Adey A., Patwardhan R. P., Qiu R., Kitzman J. O., et al. (2013) Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol 31: 1119-1125

Calabrese J. M., Sun W., Song L., Mugford J. W., Williams L., et al. (2012) Site-specific silencing of regulatory elements as a mechanism of X inactivation. Cell 151: 951-963

Capell B. C. and Berger S. L. (2013) Genome-wide epigenetics. J Invest Dermatol 133: e9

Castello A., Fischer B., Hentze M. W. and Preiss T. (2013) RNA-binding proteins in Mendelian disease. Trends Genet 29: 318-327

Cernilogar F. M., Onorati M. C., Kothe G. O., Burroughs A. M., Parsi K. M., et al. (2011) Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature 480: 391-395

Chai X., Nagarajan S., Kim K., Lee K. and Choi J. K. (2013) Regulation of the boundaries of accessible chromatin. PLoS Genet 9: e1003778

Chang C. Y., Pasolli H. A., Giannopoulou E. G., Guasch G., Gronostajski R. M., et al. (2013) NFIB is a governor of epithelial-melanocyte stem cell behaviour in a shared niche. Nature 495: 98-102

Charizanis K., Lee K. Y., Batra R., Goodwin M., Zhang C., et al. (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron 75: 437-450

Chauhan C., Zraly C. B. and Dingwall A. K. (2013) The Drosophila COMPASS-like Cmi-Trr coactivator complex regulates dpp/BMP signaling in pattern formation. Dev Biol 380: 185-198

Chaumeil J., Micsinai M., Ntziachristos P., Roth D. B., Aifantis I., et al. (2013) The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. Nat Commun 4: 2231

Chavez L., Jozefczuk J., Grimm C., Dietrich J., Timmermann B., et al. (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. Genome Res 20: 1441-1450

Chen Y., Negre N., Li Q., Mieczkowska J. O., Slattery M., et al. (2012) Systematic evaluation of factors influencing ChIP-seq fidelity. Nat Methods 9: 609-614

Chi S. W., Zang J. B., Mele A. and Darnell R. B. (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460: 479-486

Chi S. W., Hannon G. J. and Darnell R. B. (2012) An alternative mode of microRNA target recognition. Nat Struct Mol Biol 19: 321-327

Chiarle R. (2013) Translocations in normal B cells and cancers: insights from new technical approaches. Adv Immunol 117: 39-71

Chitsaz H., Yee-Greenbaum J. L., Tesler G., Lombardo M. J., Dupont C. L., et al. (2011) Efficient de novo assembly of single-cell bacterial genomes from. short-read data sets. Nat Biotechnol 29: 915-921

Cho J., Chang H., Kwon S. C., Kim B., Kim Y., et al. (2012) LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell 151: 765-777

Chopra V. S., Hendrix D. A., Core L. J., Tsui C., Lis J. T., et al. (2011) The Polycomb Group Mutant esc Leads to Augmented Levels of Paused Pol II in the Drosophila Embryo. Mol Cell 42: 837-844

Chu C., Qu K., Zhong F. L., Artandi S. E. and Chang H. Y. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 44: 667-678

Chumsakul O., Nakamura K., Kurata T., Sakamoto T., Hobman J. L., et al. (2013) High-resolution mapping of in vivo genomic transcription factor binding sites using in situ DNase I footprinting and ChIP-seq. DNA Res 20: 325-338

Churchman L. S. and Weissman J. S. (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 469: 368-373

Clouaire T., Webb S., Skene P., Illingworth R., Kerr A., et al. (2012) Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev 26: 1714-1728

Cokus S. J., Feng S., Zhang X., Chen Z., Merriman B., et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452: 215-219

Colquitt B. M., Allen W. E., Barnea G. and Lomvardas S. (2013) Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc Natl Acad Sci U S A 110: 14682-14687

Core L. J., Waterfall J. J. and Lis J. T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322: 1845-1848

Core L. J., Waterfall J. J., Gilchrist D. A., Fargo D. C., Kwak H., et al. (2012) Defining the status of RNA polymerase at promoters. Cell Rep 2: 1025-1035

Cusick M. E., Herman T. M., DePamphilis M. L. and Wassarman P. M. (1981) Structure of chromatin at deoxyribonucleic acid replication forks: prenucleosomal deoxyribonucleic acid is rapidly excised from replicating simian virus 40 chromosomes by micrococcal nuclease. Biochemistry 20: 6648-6658

Darnell J. C., Van Driesche S. J., Zhang C., Hung K. Y., Mele A., et al. (2011) FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. Cell 146: 247-261

Davis H. P. and Squire L. R. (1984) Protein synthesis and memory: a review. Psychol Bull 96: 518-559

de Wit E. and de Laat W. (2012) A decade of 3C technologies: insights into nuclear organization. Genes Dev 26: 11-24

de Wit E., Bouwman B. A., Zhu Y., Klous P., Splinter E., et al. (2013) The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature 501: 227-231

Dean F. B., Nelson J. R., Giesler T. L. and Lasken R. S. (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11: 1095-1099

Degner J. F., Pai A. A., Pique-Regi R., Veyrieras J. B., Gaffney D. J., et al. (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. Nature 482: 390-394

Dekker J., Marti-Renom M. A. and Mirny L. A. (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14: 390-403

Delpretti S., Montavon T., Leleu M., Joye E., Tzika A., et al. (2013) Multiple Enhancers Regulate Hoxd Genes and the Hotdog LncRNA during Cecum Budding. Cell Rep 5: 137-150

DeMare L. E., Leng J., Cotney J., Reilly S. K., Yin J., et al. (2013) The genomic landscape of cohesin-associated chromatin interactions. Genome Res 23: 1224-1234

Demichelis F., Setlur S. R., Banerjee S., Chakravarty D., Chen J. Y., et al. (2012) Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk. Proc Natl Acad Sci U S A 109: 6686-6691

Deng T., Zhu Z. I., Zhang S., Leng F., Cherukuri S., et al. (2013) HMGN1 modulates nucleosome occupancy and DNase I hypersensitivity at the CpG island promoters of embryonic stem cells. Mol Cell Biol 33: 3377-3389

Denholtz M., Bonora G., Chronis C., Splinter E., de Laat W., et al. (2013) Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization. Cell Stem Cell 13: 602-616

Di Ruscio A., Ebralidze A. K., Benoukraf T., Amabile G., Goff L. A., et al. (2013) DNMT1-interacting RNAs block gene-specific DNA methylation. Nature 503: 371-376

Diep D., Plongthongkum N., Gore A., Fung H. L., Shoemaker R., et al. (2012) Library-free methylation sequencing with bisulfite padlock probes. Nat Methods 9: 270-272

Dixon J. R., Selvaraj S., Yue F., Kim A., Li Y., et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485: 376-380

Dolinoy D. C., Weidman J. R., Waterland R. A. and Jirtle R. L. (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114: 567-572

Dolinoy D. C. (2008) The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev 66 Suppl 1: S7-11

Dolinoy D. C. and Faulk C. (2012) Introduction: The use of animals models to advance epigenetic science. ILAR J 53: 227-231

Dong T. G., Ho B. T., Yoder-Himes D. R. and Mekalanos J. J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. Proc Natl Acad Sci U S A 110: 2623-2628

Dostie J. and Dekker J. (2007) Mapping networks of physical interactions between genomic elements using 5C technology. Nat Protoc 2: 988-1002

Down T. A., Rakyan V. K., Turner D. J., Flicek P., Li H., et al. (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 26: 779-785

Duan Z., Andronescu M., Schutz K., Lee C., Shendure J., et al. (2012) A genome-wide 3C-method for characterizing the three-dimensional architectures of genomes. Methods 58: 277-288

Dunowska M., Biggs P. J., Zheng T. and Perrott M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418-424

Embree M., Nagarajan H., Movahedi N., Chitsaz H. and Zengler K. (2013) Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. ISME J

Farcas A. M., Blackledge N. P., Sudbery I., Long H. K., McGouran J. F., et al. (2012) KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife 1: e00205

Febrer M., McLay K., Caccamo M., Twomey K. B. and Ryan R. P. (2011) Advances in bacterial transcriptome and transposon insertion-site profiling using second-generation sequencing. Trends Biotechnol 29: 586-594

Fedoroff N. V. (2012) Presidential address. Transposable elements, epigenetics, and genome evolution. Science 338: 758-767

Feil R., Charlton J., Bird A. P., Walter J. and Reik W. (1994) Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. Nucleic Acids Res 22: 695-696

Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

Feng S., Cokus S. J., Zhang X., Chen P. Y., Bostick M., et al. (2010) Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A 107: 8689-8694

Fritsch C., Herrmann A., Nothnagel M., Szafranski K., Huse K., et al. (2012) Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res 22: 2208-2218

Frumkin D., Wasserstrom A., Kaplan S., Feige U. and Shapiro E. (2005) Genomic variability within an organism exposes its cell lineage tree. PLoS Comput Biol 1: e50

Fullwood M. J., Liu M. H., Pan Y. F., Liu J., Xu H., et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462: 58-64

Fullwood M. J., Han Y., Wei C. L., Ruan X. and Ruan Y. (2010) Chromatin interaction analysis using paired-end tag sequencing. Curr Protoc Mol Biol Chapter 21: Unit 21 15 21-25

Gaffney D. J., McVicker G., Pai A. A., Fondufe-Mittendorf Y. N., Lewellen N., et al. (2012) Controls of nucleosome positioning in the human genome. PLoS Genet 8: e1003036

Galas D. J. and Schmitz A. (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res 5: 3157-3170

Gao Q., Steine E. J., Barrasa M. I., Hockemeyer D., Pawlak M., et al. (2011) Deletion of the de novo DNA methyltransferase Dnmt3a promotes lung tumor progression. Proc Natl Acad Sci U S A 108: 18061-18066

Gehring M., Bubb K. L. and Henikoff S. (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324: 1447-1451

Gendrel A. V., Apedaile A., Coker H., Termanis A., Zvetkova I., et al. (2012) Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive x chromosome. Dev Cell 23: 265-279

Gerashchenko M. V., Lobanov A. V. and Gladyshev V. N. (2012) Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. Proc Natl Acad Sci U S A 109: 17394-17399

German M. A., Pillay M., Jeong D. H., Hetawal A., Luo S., et al. (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26: 941-946

Gertz J., Varley K. E., Reddy T. E., Bowling K. M., Pauli F., et al. (2011) Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genet 7: e1002228

Gifford W. D., Pfaff S. L. and Macfarlan T. S. (2013) Transposable elements as genetic regulatory substrates in early development. Trends Cell Biol 23: 218-226

Giresi P. G. and Lieb J. D. (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). Methods 48: 233-239

Goetz J. J. and Trimarchi J. M. (2012) Transcriptome sequencing of single cells with Smart-Seq. Nat Biotechnol 30: 763-765

Gontan C., Achame E. M., Demmers J., Barakat T. S., Rentmeester E., et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature 485: 386-390

Goodman A. L., Wu M. and Gordon J. I. (2011) Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. Nat Protoc 6: 1969-1980

Gottwein E., Corcoran D. L., Mukherjee N., Skalsky R. L., Hafner M., et al. (2011) Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines. Cell Host Microbe 10: 515-526

Grandi F. C. and An W. (2013) Non-LTR retrotransposons and microsatellites: Partners in genomic variation. Mob Genet Elements 3: e25674

Grøntved L., John S., Baek S., Liu Y., Buckley J. R., et al. (2013) C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. EMBO J 32: 1568-1583

Gschwend A. R., Weingartner L. A., Moore R. C. and Ming R. (2012) The sex-specific region of sex chromosomes in animals and plants. Chromosome Res 20: 57-69

Gu H., Smith Z. D., Bock C., Boyle P., Gnirke A., et al. (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc 6: 468-481

Gu W., Lee H. C., Chaves D., Youngman E. M., Pazour G. J., et al. (2012) CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151: 1488-1500

Guertin M. J., Martins A. L., Siepel A. and Lis J. T. (2012) Accurate prediction of inducible transcription factor binding intensities in vivo. PLoS Genet 8: e1002610

Gustems M., Woellmer A., Rothbauer U., Eck S. H., Wieland T., et al. (2013) c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. Nucleic Acids Res

Habibi E., Brinkman A. B., Arand J., Kroeze L. I., Kerstens H. H., et al. (2013) Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell 13: 360-369

Hafner M., Landgraf P., Ludwig J., Rice A., Ojo T., et al. (2008) Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods 44: 3-12

Hafner M., Landthaler M., Burger L., Khorshid M., Hausser J., et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141: 129-141

Hafner M., Landthaler M., Burger L., Khorshid M., Hausser J., et al. (2010) PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp

Hafner M., Lianoglou S., Tuschl T. and Betel D. (2012) Genome-wide identification of miRNA targets by PAR-CLIP. Methods 58: 94-105

Hah N., Danko C. G., Core L., Waterfall J. J., Siepel A., et al. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 145: 622-634

Hakim O., Sung M. H., Voss T. C., Splinter E., John S., et al. (2011) Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. Genome Res 21: 697-706

Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K. R., et al. (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484: 69-74

Hammoud S. S., Nix D. A., Zhang H., Purwar J., Carrell D. T., et al. (2009) Distinctive chromatin in human sperm packages genes for embryo development. Nature 460: 473-478

Han Y., David A., Liu B., Magadan J. G., Bennink J. R., et al. (2012) Monitoring cotranslational protein folding in mammalian cells at codon resolution. Proc Natl Acad Sci U S A 109: 12467-12472

Hancks D. C. and Kazazian H. H., Jr. (2012) Active human retrotransposons: variation and disease. Curr Opin Genet Dev 22: 191-203

Harris R. A., Wang T., Coarfa C., Nagarajan R. P., Hong C., et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol 28: 1097-1105

Hashimshony T., Wagner F., Sher N. and Yanai I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2: 666-673

He H. H., Meyer C. A., Chen M. W., Jordan V. C., Brown M., et al. (2012) Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. Genome Res 22: 1015-1025

He Y., Vogelstein B., Velculescu V. E., Papadopoulos N. and Kinzler K. W. (2008) The antisense transcriptomes of human cells. Science 322: 1855-1857

Heinz S., Romanoski C. E., Benner C., Allison K. A., Kaikkonen M. U., et al. (2013) Effect of natural genetic variation on enhancer selection and function. Nature 503: 487-492

Heisig J., Weber D., Englberger E., Winkler A., Kneitz S., et al. (2012) Target gene analysis by microarrays and chromatin immunoprecipitation identifies HEY proteins as highly redundant bHLH repressors. PLoS Genet 8: e1002728

Hiatt J. B., Pritchard C. C., Salipante S. J., O'Roak B. J. and Shendure J. (2013) Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. Genome Res 23: 843-854

Hilton I. B., Simon J. M., Lieb J. D., Davis I. J., Damania B., et al. (2013) The open chromatin landscape of Kaposi's sarcoma-associated herpesvirus. J Virol 87: 11831-11842

Hodges E., Smith A. D., Kendall J., Xuan Z., Ravi K., et al. (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res 19: 1593-1605

Hogan G. J., Lee C. K. and Lieb J. D. (2006) Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. PLoS Genet 2: e158

Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

Holwerda S. J., van de Werken H. J., Ribeiro de Almeida C., Bergen I. M., de Bruijn M. J., et al. (2013) Allelic exclusion of the immunoglobulin heavy chain locus is independent of its nuclear localization in mature B cells. Nucleic Acids Res 41: 6905-6916

Hou Y., Song L., Zhu P., Zhang B., Tao Y., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148: 873-885

Hou Y., Fan W., Yan L., Li R., Lian Y., et al. (2013) Genome analyses of single human oocytes. Cell 155: 1492-1506

Hsieh A. C., Liu Y., Edlind M. P., Ingolia N. T., Janes M. R., et al. (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature 485: 55-61

Hsieh T. F., Ibarra C. A., Silva P., Zemach A., Eshed-Williams L., et al. (2009) Genome-wide demethylation of Arabidopsis endosperm. Science 324: 1451-1454

Huang T. T., Gonzales C. B., Gu F., Hsu Y. T., Jadhav R. R., et al. (2013) Epigenetic deregulation of the anaplastic lymphoma kinase gene modulates mesenchymal characteristics of oral squamous cell carcinomas. Carcinogenesis 34: 1717-1727

Hussain S., Sajini A. A., Blanco S., Dietmann S., Lombard P., et al. (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4: 255-261

Illingworth R. S., Gruenewald-Schneider U., Webb S., Kerr A. R., James K. D., et al. (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6: e1001134

Imakaev M., Fudenberg G., McCord R. P., Naumova N., Goloborodko A., et al. (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9: 999-1003

Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223

Ingolia N. T. (2010) Genome-wide translational profiling by ribosome footprinting. Methods Enzymol 470: 119-142

Ingolia N. T., Lareau L. F. and Weissman J. S. (2011) Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell 147: 789-802

Irimia M., Tena J. J., Alexis M. S., Fernandez-Minan A., Maeso I., et al. (2012) Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. Genome Res 22: 2356-2367

Islam S., Kjallquist U., Moliner A., Zajac P., Fan J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160-1167

Islam S., Zeisel A., Joost S., La Manno G., Zajac P., et al. (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods 11: 163-166

Iwata M., Sandstrom R. S., Delrow J. J., Stamatoyannopoulos J. A. and Torok-Storb B. (2013) Functionally and Phenotypically Distinct Subpopulations of Marrow Stromal Cells Are Fibroblast in Origin and Induce Different Fates in Peripheral Blood Monocytes. Stem Cells Dev

Jacob Y., Feng S., Leblanc C. A., Bernatavichute Y. V., Stroud H., et al. (2009) ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat Struct Mol Biol 16: 763-768

Jain A., Bacolla A., Del Mundo I. M., Zhao J., Wang G., et al. (2013) DHX9 helicase is involved in preventing genomic instability induced by alternatively structured DNA in human cells. Nucleic Acids Res 41: 10345-10357

Jakopovic M., Thomas A., Balasubramaniam S., Schrump D., Giaccone G., et al. (2013) Targeting the Epigenome in Lung Cancer: Expanding Approaches to Epigenetic Therapy. Front Oncol 3: 261

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Jasmine F., Ahsan H., Andrulis I. L., John E. M., Chang-Claude J., et al. (2008) Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array. Cancer Epidemiol Biomarkers Prev 17: 3499-3508

Ji X., Zhou Y., Pandit S., Huang J., Li H., et al. (2013) SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell 153: 855-868

Jiang L., Zhang J., Wang J. J., Wang L., Zhang L., et al. (2013) Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. Cell 153: 773-784

Jiao Y. and Meyerowitz E. M. (2010) Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol 6: 419

Jin F., Li Y., Dixon J. R., Selvaraj S., Ye Z., et al. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503: 290-294

Jin S. G., Kadam S. and Pfeifer G. P. (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res 38: e125

Jullien P. E. and Berger F. (2010) DNA methylation reprogramming during plant sexual reproduction? Trends Genet 26: 394-399

Jumlongras D., Lachke S. A., O'Connell D. J., Aboukhalil A., Li X., et al. (2012) An evolutionarily conserved enhancer regulates Bmp4 expression in developing incisor and limb bud. PLoS ONE 7: e38568

Jungkamp A. C., Stoeckius M., Mecenas D., Grun D., Mastrobuoni G., et al. (2011) In vivo and transcriptome-wide identification of RNA binding protein target sites. Mol Cell 44: 828-840

Kadkhodaei B., Alvarsson A., Schintu N., Ramsköld D., Volakakis N., et al. (2013) Transcription factor Nurr1 maintains fiber integrity and nuclear-encoded mitochondrial gene expression in dopamine neurons. Proc Natl Acad Sci U S A 110: 2360-2365

Kahrstrom C. T. (2013) Epigenetics: Legionella makes its mark on histones. Nat Rev Genet 14: 370

Kaikkonen M. U., Spann N. J., Heinz S., Romanoski C. E., Allison K. A., et al. (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol Cell 51: 310-325

Kalisky T. and Quake S. R. (2011) Single-cell genomics. Nat Methods 8: 311-314

Kaneko S., Bonasio R., Saldana-Meyer R., Yoshida T., Son J., et al. (2014) Interactions between JARID2 and Noncoding RNAs Regulate PRC2 Recruitment to Chromatin. Mol Cell 53: 290-300

Kanematsu S., Tanimoto K., Suzuki Y. and Sugano S. (2014) Screening for possible miRNA-mRNA associations in a colon cancer cell line. Gene 533: 520-531

Kennedy S. R., Salk J. J., Schmitt M. W. and Loeb L. A. (2013) Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. PLoS Genet 9: e1003794

Kent N. A., Adams S., Moorhouse A. and Paszkiewicz K. (2011) Chromatin particle spectrum analysis: a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. Nucleic Acids Res 39: e26

Kim M., Park Y. K., Kang T. W., Lee S. H., Rhee Y. H., et al. (2013) Dynamic changes in DNA methylation and hydroxymethylation when hES cells undergo differentiation toward a neuronal lineage. Hum Mol Genet 23: 657-667

Kim Y. J., Greer C. B., Cecchini K. R., Harris L. N., Tuck D. P., et al. (2013) HDAC inhibitors induce transcriptional repression of high copy number genes in breast cancer through elongation blockade. Oncogene 32: 2828-2835

Kishore S., Jaskiewicz L., Burger L., Hausser J., Khorshid M., et al. (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods 8: 559-564

Kivioja T., Vaharautio A., Karlsson K., Bonke M., Enge M., et al. (2012) Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 9: 72-74

Klein I. A., Resch W., Jankovic M., Oliveira T., Yamane A., et al. (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147: 95-106

Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

Kobayashi H., Sakurai T., Imai M., Takahashi N., Fukuda A., et al. (2012) Contribution of Intragenic DNA Methylation in Mouse Gametic DNA Methylomes to Establish Oocyte-Specific Heritable Marks. PLoS Genet 8: e1002440

Kobayashi H., Sakurai T., Miura F., Imai M., Mochiduki K., et al. (2013) High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. Genome Res 23: 616-627

Koh K. P. and Rao A. (2013) DNA methylation and methylcytosine oxidation in cell fate decisions. Curr Opin Cell Biol 25: 152-16

Koldamova R., Schug J., Lefterova M., Cronican A. A., Fitz N. F., et al. (2014) Genome-wide approaches reveal EGR1-controlled regulatory networks associated with neurodegeneration. Neurobiol Dis 63: 107-114

Konig J., Zarnack K., Rot G., Curk T., Kayikci M., et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17: 909-915

Kovalchuk A. L., Ansarah-Sobrinho C., Hakim O., Resch W., Tolarova H., et al. (2012) Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. Proc Natl Acad Sci U S A 109: 10972-10977

Kozlenkov A., Roussos P., Timashpolsky A., Barbu M., Rudchenko S., et al. (2014) Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. Nucleic Acids Res 42: 109-127

Kudla G., Granneman S., Hahn D., Beggs J. D. and Tollervey D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci U S A 108: 10010-10015

Kumar V., Muratani M., Rayan N. A., Kraus P., Lufkin T., et al. (2013) Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol 31: 615-622

Kuo T., Lew M. J., Mayba O., Harris C. A., Speed T. P., et al. (2012) Genome-wide analysis of glucocorticoid receptor-binding sites in myotubes identifies gene networks modulating insulin signaling. Proc Natl Acad Sci U S A 109: 11160-11165

Lai C. F., Flach K. D., Alexi X., Fox S. P., Ottaviani S., et al. (2013) Co-regulated gene expression by oestrogen receptor alpha and liver receptor homolog-1 is a feature of the oestrogen response in breast cancer cells. Nucleic Acids Res 41: 10228-10240

Lai D., Proctor J. R. and Meyer I. M. (2013) On the importance of cotranscriptional RNA structure formation. RNA 19: 1461-1473

Lam M. T., Cho H., Lesch H. P., Gosselin D., Heinz S., et al. (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498: 511-515

Lan X., Witt H., Katsumura K., Ye Z., Wang Q., et al. (2012) Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. Nucleic Acids Res 40: 7690-7704

Larschan E., Bishop E. P., Kharchenko P. V., Core L. J., Lis J. T., et al. (2011) X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature 471: 115-118

Larson M. H., Gilbert L. A., Wang X., Lim W. A., Weissman J. S., et al. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc 8: 2180-2196

Lasken R. S. (2013) Single-cell sequencing in its prime. Nat Biotechnol 31: 211-212

Lassen K. S., Schultz H., Heegaard N. H. and He M. (2012) A novel DNAseq program for enhanced analysis of Illumina GAII data: a case study on antibody complementarity-determining regions. N Biotechnol 29: 271-278

Law G. L., Korth M. J., Benecke A. G. and Katze M. G. (2013) Systems virology: host-directed approaches to viral pathogenesis and drug targeting. Nat Rev Microbiol 11: 455-466

Lazarovici A., Zhou T., Shafer A., Dantas Machado A. C., Riley T. R., et al. (2013) Probing DNA shape and methylation state on a genomic scale with DNase I. Proc Natl Acad Sci U S A 110: 6376-6381

Lebedeva S., Jens M., Theil K., Schwanhausser B., Selbach M., et al. (2011) Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. Mol Cell 43: 340-352

Lee J., Seok S., Yu P., Kim K., Smith Z., et al. (2012) Genomic analysis of hepatic farnesoid X receptor binding sites reveals altered binding in obesity and direct gene repression by farnesoid X receptor in mice. Hepatology 56: 108-117

Lee M. T., Bonneau A. R., Takacs C. M., Bazzini A. A., DiVito K. R., et al. (2013) Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature 503: 360-364

Lee S., Liu B., Lee S., Huang S. X., Shen B., et al. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A 109: E2424-2432

Lee T. I. and Young R. A. (2013) Transcriptional regulation and its misregulation in disease. Cell 152: 1237-1251

Lesch B. J., Dokshin G. A., Young R. A., McCarrey J. R. and Page D. C. (2013) A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. Proc Natl Acad Sci U S A 110: 16061-16066

Li G., Fullwood M. J., Xu H., Mulawadi F. H., Velkov S., et al. (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol 11: R22

Li G., Ruan X., Auerbach R. K., Sandhu K. S., Zheng M., et al. (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148: 84-98

Li G. W., Oh E. and Weissman J. S. (2012) The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature 484: 538-541

Li N., Ye M., Li Y., Yan Z., Butcher L. M., et al. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. Methods 52: 203-212

Li P., Spann N. J., Kaikkonen M. U., Lu M., Oh da Y., et al. (2013) NCoR repression of LXRs restricts macrophage biosynthesis of insulin-sensitizing omega 3 fatty acids. Cell 155: 200-214

Li W., Notani D., Ma Q., Tanasa B., Nunez E., et al. (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498: 516-520

Li Z., Chao T. C., Chang K. Y., Lin N., Patil V. S., et al. (2014) The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. Proc Natl Acad Sci U S A 111: 1002-1007

R. B. (2010) RNA processing and its regulation: global insights into biological networks. Nat Rev Genet 11: 75-87

Lieberman-Aiden E., van Berkum N. L., Williams L., Imakaev M., Ragoczy T., et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289-293

Lister R., Mukamel E. A., Nery J. R., Urich M., Puddifoot C. A., et al. (2013) Global epigenomic reconfiguration during mammalian brain development. Science 341: 1237905

Liu B., Han Y. and Qian S. B. (2013) Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. Mol Cell 49: 453-463

Liu M., Li C. L., Stamatoyannopoulos G., Dorschner M. O., Humbert R., et al. (2012) Gammaretroviral Vector Integration Occurs Overwhelmingly Within and Near DNase Hypersensitive Sites. Hum Gene Ther 23: 231-237

Liu X., Jiang H., Gu Z. and Roberts J. W. (2013) High-resolution view of bacteriophage lambda gene expression by ribosome profiling. Proc Natl Acad Sci U S A 110: 11928-11933

Liu Y., Hu W., Murakawa Y., Yin J., Wang G., et al. (2013) Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. Sci Rep 3: 2054

Lo K. A., Labadorf A., Kennedy N. J., Han M. S., Yap Y. S., et al. (2013) Analysis of in vitro insulin-resistance models and their physiological relevance to in vivo diet-induced adipose insulin resistance. Cell Rep 5: 259-270

Lovett M. (2013) The applications of single-cell genomics. Hum Mol Genet 22: R22-26

Lucks J. B., Mortimer S. A., Trapnell C., Luo S., Aviran S., et al. (2011) Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Proc Natl Acad Sci U S A 108: 11063-11068

Lun F. M., Chiu R. W., Sun K., Leung T. Y., Jiang P., et al. (2013) Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. Clin Chem 59: 1583-1594

Lyko F., Foret S., Kucharski R., Wolf S., Falckenhayn C., et al. (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol 8: e1000506

Macosko E. Z. and McCarroll S. A. (2013) Genetics. Our fallen genomes. Science 342: 564-565

Maehara K., Odawara J., Harada A., Yoshimi T., Nagao K., et al. (2013) A co-localization model of paired ChIP-seq data using a large ENCODE data set enables comparison of multiple samples. Nucleic Acids Res 41: 54-62

Majoros W. H., Lekprasert P., Mukherjee N., Skalsky R. L., Corcoran D. L., et al. (2013) MicroRNA target site identification by integrating sequence and binding information. Nat Methods 10: 630-633

Malhotra A., Lindberg M., Faust G. G., Leibowitz M. L., Clark R. A., et al. (2013) Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. Genome Res 23: 762-776

Mandal P. K., Ewing A. D., Hancks D. C. and Kazazian H. H., Jr. (2013) Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. Hum Mol Genet 22: 3730-3748

Mann B., van Opijnen T., Wang J., Obert C., Wang Y. D., et al. (2012) Control of virulence by small RNAs in Streptococcus pneumoniae. PLoS Pathog 8: e1002788

Marinov G. K., Williams B. A., McCue K., Schroth G. P., Gertz J., et al. (2014) From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. Genome Res 24: 496-510

Maruyama H., Harwood J. C., Moore K. M., Paszkiewicz K., Durley S. C., et al. (2013) An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis. EMBO Rep 14: 711-717

Marwick J. A., Kirkham P. A., Stevenson C. S., Danahay H., Giddings J., et al. (2004) Cigarette smoke alters chromatin remodeling and induces proinflammatory genes in rat lungs. Am J Respir Cell Mol Biol 31: 633-642

Matsushita H., Vesely M. D., Koboldt D. C., Rickert C. G., Uppaluri R., et al. (2012) Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature 482: 400-404

Maunakea A. K., Nagarajan R. P., Bilenky M., Ballinger T. J., D'Souza C., et al. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466: 253-257

Maurano M. T., Humbert R., Rynes E., Thurman R. E., Haugen E., et al. (2012) Systematic localization of common disease-associated variation in regulatory DNA. Science 337: 1190-1195

Mazzoni E. O., Mahony S., Closser M., Morrison C. A., Nedelec S., et al. (2013) Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. Nat Neurosci 16: 1219-1227

McConnell M. J., Lindberg M. R., Brennand K. J., Piper J. C., Voet T., et al. (2013) Mosaic copy number variation in human neurons. Science 342: 632-637

McDaniell R., Lee B. K., Song L., Liu Z., Boyle A. P., et al. (2010) Heritable individual-specific and allele-specific chromatin signatures in humans. Science 328: 235-239

McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

McKenna L. B., Schug J., Vourekas A., McKenna J. B., Bramswig N. C., et al. (2010) MicroRNAs control intestinal epithelial differentiation, architecture, and barrier function. Gastroenterology 139: 1654-1664, 1664 e1651

McLean J. S., Lombardo M. J., Ziegler M. G., Novotny M., Yee-Greenbaum J., et al. (2013) Genome of the pathogen Porphyromonas gingivalis recovered from a biofilm in a hospital sink using a high-throughput single-cell genomics platform. Genome Res 23: 867-877

McMullen P. D., Bhattacharya S., Woods C. G., Sun B., Yarborough K., et al. (2014) A map of the PPARalpha transcription regulatory network for primary human hepatocytes. Chem Biol Interact 209: 14-24

Meaburn E. and Schulz R. (2012) Next generation sequencing in epigenetics: insights and challenges. Semin Cell Dev Biol 23: 192-199

Medvedovic J., Ebert A., Tagoh H., Tamir I. M., Schwickert T. A., et al. (2013) Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. Immunity 39: 229-244

Meissner A., Gnirke A., Bell G. W., Ramsahoye B., Lander E. S., et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33: 5868-5877

Meissner A., Mikkelsen T. S., Gu H., Wernig M., Hanna J., et al. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454: 766-770

Mellen M., Ayata P., Dewell S., Kriaucionis S. and Heintz N. (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151: 1417-1430

Mendoza-Parra M. A., Van Gool W., Mohamed Saleem M. A., Ceschin D. G. and Gronemeyer H. (2013) A quality control system for profiles obtained by ChIP sequencing. Nucleic Acids Res 41: e196

Meredith D. M., Borromeo M. D., Deering T. G., Casey B. H., Savage T. K., et al. (2013) Program specificity for Ptf1a in pancreas versus neural tube development correlates with distinct collaborating cofactors and chromatin accessibility. Mol Cell Biol 33: 3166-3179

Meyer K. D., Saletore Y., Zumbo P., Elemento O., Mason C. E., et al. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149: 1635-1646

Mitchell J. A., Clay I., Umlauf D., Chen C. Y., Moir C. A., et al. (2012) Nuclear RNA sequencing of the mouse erythroid cell transcriptome. PLoS One 7: e49274

Miura F., Enomoto Y., Dairiki R. and Ito T. (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res 40: e136

Mukherjee N., Corcoran D. L., Nusbaum J. D., Reid D. W., Georgiev S., et al. (2011) Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. Mol Cell 43: 327-339

Murray S. M., Panis G., Fumeaux C., Viollier P. H. and Howard M. (2013) Computational and genetic reduction of a cell cycle to its simplest, primordial components. PLoS Biol 11: e1001749

Murtaza M., Dawson S. J., Tsui D. W., Gale D., Forshew T., et al. (2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 497: 108-112

Myllykangas S., Buenrostro J. D., Natsoulis G., Bell J. M. and Ji H. P. (2011) Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing. Nat Biotechnol 29: 1024-1027

Nagano T., Lubling Y., Stevens T. J., Schoenfelder S., Yaffe E., et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502: 59-64

Nagarajavel V., Iben J. R., Howard B. H., Maraia R. J. and Clark D. J. (2013) Global 'bootprinting' reveals the elastic architecture of the yeast TFIIIB-TFIIIC transcription complex in vivo. Nucleic Acids Res 41: 8135-8143

Navin N. and Hicks J. (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3: 31

Navin N., Kendall J., Troge J., Andrews P., Rodgers L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90-94

Neph S., Vierstra J., Stergachis A. B., Reynolds A. P., Haugen E., et al. (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 489: 83-90

Neri F., Krepelova A., Incarnato D., Maldotti M., Parlato C., et al. (2013) Dnmt3L Antagonizes DNA Methylation at Bivalent Promoters and Favors DNA Methylation at Gene Bodies in ESCs. Cell 155: 121-134

Neyret-Kahn H., Benhamed M., Ye T., Le Gras S., Cossec J. C., et al. (2013) Sumoylation at chromatin governs coordinated repression of a transcriptional program essential for cell growth and proliferation. Genome Res 23: 1563-1579

Ni T., Corcoran D. L., Rach E. A., Song S., Spana E. P., et al. (2010) A paired-end sequencing strategy to map the complex landscape of transcription initiation. Nat Methods 7: 521-527

Ni X., Zhuo M., Su Z., Duan J., Gao Y., et al. (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110: 21083-21088

Nishida H., Katayama T., Suzuki Y., Kondo S. and Horiuchi H. (2013) Base composition and nucleosome density in exonic and intronic regions in genes of the filamentous ascomycetes Aspergillus nidulans and Aspergillus oryzae. Gene 525: 10-May

Nora E. P., Lajoie B. R., Schulz E. G., Giorgetti L., Okamoto I., et al. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485: 381-385

O'Huallachain M., Weissman S. M. and Snyder M. P. (2013) The variable somatic genome. Cell Cycle 12: 5-6

Oh E., Becker A. H., Sandikci A., Huber D., Chaba R., et al. (2011) Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell 147: 1295-1308

Oliveira T., Resch W., Jankovic M., Casellas R., Nussenzweig M. C., et al. (2011) Translocation capture sequencing: A method for high throughput mapping of chromosomal rearrangements. J Immunol Methods 375: 176-181

Oliveira T. Y., Resch W., Jankovic M., Casellas R., Nussenzweig M. C., et al. (2012) Translocation capture sequencing: a method for high throughput mapping of chromosomal rearrangements. J Immunol Methods 375: 176-181

Olovnikov I., Ryazansky S., Shpiz S., Lavrov S., Abramov Y., et al. (2013) De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. Nucleic Acids Res 41: 5757-5768

Osborne R. J. and Thornton C. A. (2006) RNA-dominant diseases. Hum Mol Genet 15 Spec No 2: R162-169

Paul D. S., Nisbet J. P., Yang T. P., Meacham S., Rendon A., et al. (2011) Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. PLoS Genet 7: e1002139

Paul D. S., Albers C. A., Rendon A., Voss K., Stephens J., et al. (2013) Maps of open chromatin highlight cell type-restricted patterns of regulatory sequence variation at hematological trait loci. Genome Res 23: 1130-1141

Pelechano V., Wei W. and Steinmetz L. M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature 497: 127-131

Pfeifer G. P., Kadam S. and Jin S. G. (2013) 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin 6: 10

Piccolo F. M. and Fisher A. G. (2014) Getting rid of DNA methylation. Trends Cell Biol 24: 136-143

Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S. Winberg G., et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10: 1096-1098

Pinello L., Xu J., Orkin S. H. and Yuan G. C. (2014) Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. Proc Natl Acad Sci U S A 111: E344-353

Polymenidou M., Lagier-Tourenne C., Hutt K. R., Huelga S. C., Moran J., et al. (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci 14: 459-468

Pomraning K. R., Smith K. M. and Freitag M. (2009) Genome-wide high throughput analysis of DNA methylation in eukaryotes. Methods 47: 142-150

Ponts N., Harris E. Y., Prudhomme J., Wick I., Eckhardt-Ludka C., et al. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res 20: 228-238

Potok M. E., Nix D. A., Parnell T. J. and Cairns B. R. (2013) Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. Cell 153: 759-772

Potter N. E., Ermini L., Papaemmanuil E., Cazzaniga G., Vijayaraghavan G., et al. (2013) Single-cell mutational profiling and clonal phylogeny in cancer. Genome Res 23: 2115-2125

Poulos M. G., Batra R., Li M., Yuan Y., Zhang C., et al. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. Hum Mol Genet 22: 3547-3558

Prickett A. R., Barkas N., McCole R. B., Hughes S., Amante S. M., et al. (2013) Genome-wide and parental allele-specific analysis of CTCF and cohesin DNA binding in mouse brain reveals a tissue-specific binding pattern and an association with imprinted differentially methylated regions. Genome Res 23: 1624-1635

Puszyk W., Down T., Grimwade D., Chomienne C., Oakey R. J., et al. (2013) The epigenetic regulator PLZF represses L1 retrotransposition in germ and progenitor cells. EMBO J 32: 1941-1952

Qi L. S., Larson M. H., Gilbert L. A., Doudna J. A., Weissman J. S., et al. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152: 1173-1183

Qi X., Daily K., Nguyen K., Wang H., Mayhew D., et al. (2012) Retrotransposon profiling of RNA polymerase III initiation sites. Genome Res 22: 681-692

Ramasubramanyan S., Kanhere A., Osborn K., Flower K., Jenner R. G., et al. (2012) Genome-wide analyses of Zta binding to the Epstein-Barr virus genome reveals interactions in both early and late lytic cycles and an epigenetic switch leading to an altered binding profile. J Virol 86: 12494-12502

Ramskold D., Luo S., Wang Y. C., Li R., Deng Q., et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30: 777-782

Rauch T. A., Zhong X., Wu X., Wang M., Kernstine K. H., et al. (2008) High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U S A 105: 252-257

Rauch T. A. and Pfeifer G. P. (2009) The MIRA method for DNA methylation analysis. Methods Mol Biol 507: 65-75

Regulski M., Lu Z., Kendall J., Donoghue M. T., Reinders J., et al. (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res 23: 1651-1662

Rich A. and RajBhandary U. L. (1976) Transfer RNA: molecular structure, sequence, and properties. Annu Rev Biochem 45: 805-860

Riley K. J., Rabinowitz G. S., Yario T. A., Luna J. M., Darnell R. B., et al. (2012) EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. EMBO J 31: 2207-2221

Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

Robyr D., Friedli M., Gehrig C., Arcangeli M., Marin M., et al. (2011) Chromosome conformation capture uncovers potential genome-wide interactions between human conserved non-coding sequences. PLoS ONE 6: e17634

Rocha P. P., Micsinai M., Kim J. R., Hewitt S. L., Souza P. P., et al. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47: 873-885

Rocha P. P., Chaumeil J. and Skok J. A. (2013) Molecular biology. Finding the right partner in a 3D genome. Science 342: 1333-1334

Rodley C. D., Pai D. A., Mills T. A., Engelke D. R. and O'Sullivan J. M. (2011) tRNA gene identity affects nuclear positioning. PLoS One 6: e29267

Rodrigues J. A., Ruan R., Nishimura T., Sharma M. K., Sharma R., et al. (2013) Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. Proc Natl Acad Sci U S A 110: 7934-7939

Rodriguez B. A., Frankhouser D., Murphy M., Trimarchi M., Tam H. H., et al. (2012) Methods for high-throughput MethylCap-Seq data analysis. BMC Genomics 13 Suppl 6: S14

Rogelj B., Easton L. E., Bogu G. K., Stanton L. W., Rot G., et al. (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci Rep 2: 603

Ronan J. L., Wu W. and Crabtree G. R. (2013) From neural development to cognition: unexpected roles for chromatin. Nat Rev Genet 14: 347-359

Rosseel T., Scheuch M., Hoper D., De Regge N., Caij A. B., et al. (2012) DNase SISPA-next generation sequencing confirms Schmallenberg virus in Belgian field samples and identifies genetic variation in Europe. PLoS One 7: e41967

Ruike Y., Imanaka Y., Sato F., Shimizu K. and Tsujimoto G. (2010) Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. BMC Genomics 11: 137

Saied M. H., Marzec J., Khalid S., Smith P., Down T. A., et al. (2012) Genome wide analysis of acute myeloid leukemia reveal leukemia specific methylome and subtype specific hypomethylation of repeats. PLoS One 7: e33213

Sajan S. A. and Hawkins R. D. (2012) Methods for identifying higher-order chromatin structure. Annu Rev Genomics Hum Genet 13: 59-82

Sakabe N. J., Aneas I., Shen T., Shokri L., Park S. Y., et al. (2012) Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function. Hum Mol Genet 21: 2194-2204

Sakurai M., Yano T., Kawabata H., Ueda H. and Suzuki T. (2010) Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. Nat Chem Biol 6: 733-740

Salton M., Elkon R., Borodina T., Davydov A., Yaspo M. L., et al. (2011) Matrin 3 binds and stabilizes mRNA. PLoS One 6: e23882

Sanyal A., Lajoie B. R., Jain G. and Dekker J. (2012) The long-range interaction landscape of gene promoters. Nature 489: 109-113

Sarmiento F., Mrázek J. and Whitman W. B. (2013) Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon Methanococcus maripaludis. Proc Natl Acad Sci U S A 110: 4726-4731

Sasagawa Y., Nikaido I., Hayashi T., Danno H., Uno K. D., et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14: R31

Sati S., Tanwar V. S., Kumar K. A., Patowary A., Jain V., et al. (2012) High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. PLoS One 7: e31621

Saunders A., Core L. J., Sutcliffe C., Lis J. T. and Ashe H. L. (2013) Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. Genes Dev 27: 1146-1158

Schauer T., Schwalie P. C., Handley A., Margulies C. E., Flicek P., et al. (2013) CAST-ChIP maps cell-type-specific chromatin states in the Drosophila central nervous system. Cell Rep 5: 271-282

Schillebeeckx M., Schrade A., Lobs A. K., Pihlajoki M., Wilson D. B., et al. (2013) Laser capture microdissection-reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse. Nucleic Acids Res 41: e116

Schlesinger F., Smith A. D., Gingeras T. R., Hannon G. J. and Hodges E. (2013) De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements. Genome Res 23: 1601-1614

Schmitt M. W., Kennedy S. R., Salk J. J., Fox E. J., Hiatt J. B., et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci U S A 109: 14508-14513

Schmitz R. J., He Y., Valdes-Lopez O., Khan S. M., Joshi T., et al. (2013) Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res 23: 1663-1674

Schmolka N., Serre K., Grosso A. R., Rei M., Pennington D. J., et al. (2013) Epigenetic and transcriptional signatures of stable versus plastic differentiation of proinflammatory gammadelta T cell subsets. Nat Immunol 14: 1093-1100

Schones D. E., Cui K., Cuddapah S., Roh T. Y., Barski A., et al. (2008) Dynamic regulation of nucleosome positioning in the human genome. Cell 132: 887-898

Schweiger M. R., Barmeyer C. and Timmermann B. (2013) Genomics and epigenomics: new promises of personalized medicine for cancer patients. Brief Funct Genomics 12: 411-421

Seisenberger S., Andrews S., Krueger F., Arand J., Walter J., et al. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell 48: 849-862

Serre D., Lee B. H. and Ting A. H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res 38: 391-399

Seth-Smith H. M., Harris S. R., Skilton R. J., Radebe F. M., Golparian D., et al. (2013) Whole-genome sequences of Chlamydia trachomatis directly from clinical samples without culture. Genome Res 23: 855-866

Sexton T., Kurukuti S., Mitchell J. A., Umlauf D., Nagano T., et al. (2012) Sensitive detection of chromatin coassociations using enhanced chromosome conformation capture on chip. Nat Protoc 7: 1335-1350

Shalek A. K., Satija R., Adiconis X., Gertner R. S., Gaublomme J. T., et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498: 236-240

Shang W. H., Hori T., Martins N. M., Toyoda A., Misu S., et al. (2013) Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. Dev Cell 24: 635-648

Shapiro E., Biezuner T. and Linnarsson S. (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14: 618-630

Shen H., Qiu C., Li J., Tian Q. and Deng H. W. (2013) Characterization of the DNA methylome and its interindividual variation in human peripheral blood monocytes. Epigenomics 5: 255-269

Shirane K., Toh H., Kobayashi H., Miura F., Chiba H., et al. (2013) Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet 9: e1003439

Shiroguchi K., Jia T. Z., Sims P. A. and Xie X. S. (2012) Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. Proc Natl Acad Sci U S A 109: 1347-1352

Shukla R., Upton K. R., Munoz-Lopez M., Gerhardt D. J., Fisher M. E., et al. (2013) Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell 153: 101-111

Sievers C., Schlumpf T., Sawarkar R., Comoglio F. and Paro R. (2012) Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. Nucleic Acids Res 40: e160

Simon J. M., Giresi P. G., Davis I. J. and Lieb J. D. (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc 7: 256-267

Skalsky R. L., Corcoran D. L., Gottwein E., Frank C. L., Kang D., et al. (2012) The viral and cellular microRNA targetome in lymphoblastoid cell lines. PLoS Pathog 8: e1002484

Skurnik D., Roux D., Cattoir V., Danilchanka O., Lu X., et al. (2013) Enhanced in vivo fitness of carbapenem-resistant oprD mutants of Pseudomonas aeruginosa revealed through high-throughput sequencing. Proc Natl Acad Sci U S A 110: 20747-20752

Smallwood S. A., Tomizawa S., Krueger F., Ruf N., Carli N., et al. (2011) Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat Genet 43: 811-814

Smith Z. D., Gu H., Bock C., Gnirke A. and Meissner A. (2009) High-throughput bisulfite sequencing in mammalian genomes. Methods 48: 226-232

Smith Z. D. and Meissner A. (2013) DNA methylation: roles in mammalian development. Nat Rev Genet 14: 204-220

Soleimani V. D., Palidwor G. A., Ramachandran P., Perkins T. J. and Rudnicki M. A. (2013) Chromatin tandem affinity purification sequencing. Nat Protoc 8: 1525-1534

Solomon M. J., Larsen P. L. and Varshavsky A. (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell 53: 937-947

Solyom S., Ewing A. D., Rahrmann E. P., Doucet T., Nelson H. H., et al. (2012) Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res 22: 2328-2338

Song C. X., Szulwach K. E., Dai Q., Fu Y., Mao S. Q., et al. (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell 153: 678-691

Song L., Zhang Z., Grasfeder L. L., Boyle A. P., Giresi P. G., et al. (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res 21: 1757-1767

Spaethling J. M. and Eberwine J. H. (2013) Single-cell transcriptomics for drug target discovery. Curr Opin Pharmacol 13: 786-790

Splinter E., de Wit E., van de Werken H. J., Klous P. and de Laat W. (2012) Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods 58: 221-230

Stadhouders R., Kolovos P., Brouwer R., Zuin J., van den Heuvel A., et al. (2013) Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. Nat Protoc 8: 509-524

Stadler M., Artiles K., Pak J. and Fire A. (2012) Contributions of mRNA abundance, ribosome loading, and post- or peri-translational effects to temporal repression of C. elegans heterochronic miRNA targets. Genome Res 22: 2418-2426

Stadler M. B., Murr R., Burger L., Ivanek R., Lienert F., et al. (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480: 490-495

Stevens M., Cheng J. B., Li D., Xie M., Hong C., et al. (2013) Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. Genome Res 23: 1541-1553

Stevens M., Cheng J. B., Li D., Xie M., Hong C., et al. (2013) Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. Genome Res 23: 1541-1553

Stoll G., Pietilainen O. P., Linder B., Suvisaari J., Brosi C., et al. (2013) Deletion of TOP3beta, a component of FMRP-containing mRNPs, contributes to neurodevelopmental disorders. Nat Neurosci 16: 1228-1237

Stumpf C. R., Moreno M. V., Olshen A. B., Taylor B. S. and Ruggero D. (2013) The translational landscape of the Mammalian cell cycle. Mol Cell 52: 574-582

Taiwo O., Wilson G. A., Morris T., Seisenberger S., Reik W., et al. (2012) Methylome analysis using MeDIP-seq with low DNA concentrations. Nat Protoc 7: 617-636

Tan L., Xiong L., Xu W., Wu F., Huang N., et al. (2013) Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. Nucleic Acids Res 41: e84

Tan S. K., Lin Z. H., Chang C. W., Varang V., Chng K. R., et al. (2011) AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. EMBO J 30: 2569-2581

Tanaka Y., Joshi A., Wilson N. K., Kinston S., Nishikawa S., et al. (2012) The transcriptional programme controlled by Runx1 during early embryonic blood development. Dev Biol 366: 404-419

Tang B., Becanovic K., Desplats P. A., Spencer B., Hill A. M., et al. (2012) Forkhead box protein p1 is a transcriptional repressor of immune signaling in the CNS: implications for transcriptional dysregulation in Huntington disease. Hum Mol Genet 21: 3097-3111

Tang D. T., Plessy C., Salimullah M., Suzuki A. M., Calligaris R., et al. (2013) Suppression of artifacts and barcode bias in high-throughput transcriptome analyses utilizing template switching. Nucleic Acids Res 41: e44

Telese F., Gamliel A., Skowronska-Krawczyk D., Garcia-Bassets I. and Rosenfeld M. G. (2013) “Seq-ing” insights into the epigenetics of neuronal gene regulation. Neuron 77: 606-623

Thapar R., Denmon A. P. and Nikonowicz E. P. (2014) Recognition modes of RNA tetraloops and tetraloop-like motifs by RNA-binding proteins. Wiley Interdiscip Rev RNA 5: 49-67

Thomson J. P., Lempiainen H., Hackett J. A., Nestor C. E., Muller A., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. Genome Biol 13: R93

Thomson J. P., Hunter J. M., Nestor C. E., Dunican D. S., Terranova R., et al. (2013) Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. Nucleic Acids Res 41: e206-

Tollervey J. R., Curk T., Rogelj B., Briese M., Cereda M., et al. (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat Neurosci 14: 452-458

Tolstorukov M. Y., Sansam C. G., Lu P., Koellhoffer E. C., Helming K. C., et al. (2013) Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. Proc Natl Acad Sci U S A 110: 10165-10170

Troy E. B., Lin T., Gao L., Lazinski D. W., Camilli A., et al. (2013) Understanding barriers to Borrelia burgdorferi dissemination during infection using massively parallel sequencing. Infect Immun 81: 2347-2357

Turnbaugh P. J., Ridaura V. K., Faith J. J., Rey F. E., Knight R., et al. (2009) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1: 6ra14

Turnbaugh P. J., Quince C., Faith J. J., McHardy A. C., Yatsunenko T., et al. (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. Proc Natl Acad Sci U S A 107: 7503-7508

Turner J. R., Ray R., Lee B., Everett L., Xiang J., et al. (2013) Evidence from mouse and man for a role of neuregulin 3 in nicotine dependence. Mol Psychiatry

Twigg S. R., Vorgia E., McGowan S. J., Peraki I., Fenwick A. L., et al. (2013) Reduced dosage of ERF causes complex craniosynostosis in humans and mice and links ERK1/2 signaling to regulation of osteogenesis. Nat Genet 45: 308-313

Udan-Johns M., Bengoechea R., Bell S., Shao J., Diamond M. I., et al. (2014) Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. Hum Mol Genet 23: 157-170

Underwood J. G., Uzilov A. V., Katzman S., Onodera C. S., Mainzer J. E., et al. (2010) FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat Methods 7: 995-1001

Uniacke J., Holterman C. E., Lachance G., Franovic A., Jacob M. D., et al. (2012) An oxygen-regulated switch in the protein synthesis machinery. Nature 486: 126-129

Valentim C. L., LoVerde P. T., Anderson T. J. and Criscione C. D. (2009) Efficient genotyping of Schistosoma mansoni miracidia following whole genome amplification. Mol Biochem Parasitol 166: 81-84

van de Werken H. J., Landan G., Holwerda S. J., Hoichman M., Klous P., et al. (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9: 969-972

Van der Aa N., Esteki M. Z., Vermeesch J. R. and Voet T. (2013) Preimplantation genetic diagnosis guided by single-cell genomics. Genome Med 5: 71

van Opijnen T., Bodi K. L. and Camilli A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 6: 767-772

van Opijnen T. and Camilli A. (2012) A fine scale phenotype-genotype virulence map of a bacterial pathogen. Genome Res 22: 2541-2551

van Opijnen T. and Camilli A. (2013) Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol 11: 435-442

Veluchamy A., Lin X., Maumus F., Rivarola M., Bhavsar J., et al. (2013) Insights into the role of DNA methylation in diatoms by genome-wide profiling in Phaeodactylum tricornutum. Nat Commun 4: 2091

Verma-Gaur J., Torkamani A., Schaffer L., Head S. R., Schork N. J., et al. (2012) Noncoding transcription within the Igh distal V(H) region at PAIR elements affects the 3D structure of the Igh locus in pro-B cells. Proc Natl Acad Sci U S A 109: 17004-17009

Vidaki A., Daniel B. and Court D. S. (2013) Forensic DNA methylation profiling--Potential opportunities and challenges. Forensic Sci Int Genet 7: 499-507

Voet T., Kumar P., Van Loo P., Cooke S. L., Marshall J., et al. (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res 41: 6119-6138

Vourekas A., Zheng Q., Alexiou P., Maragkakis M., Kirino Y., et al. (2012) Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat Struct Mol Biol 19: 773-781

Wan Y., Qu K., Ouyang Z. and Chang H. Y. (2013) Genome-wide mapping of RNA structure using nuclease digestion and high-throughput sequencing. Nat Protoc 8: 849-869

Wang D., Garcia-Bassets I., Benner C., Li W., Su X., et al. (2011) Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474: 390-394

Wang Q., Gu L., Adey A., Radlwimmer B., Wang W., et al. (2013) Tagmentation-based whole-genome bisulfite sequencing. Nat Protoc 8: 2022-2032

Wang T., Cui Y., Jin J., Guo J., Wang G., et al. (2013) Translating mRNAs strongly correlate to proteins in a multivariate manner and their translation ratios are phenotype specific. Nucleic Acids Res 41: 4743-4754

Wang T., Wu H., Li Y., Szulwach K. E., Lin L., et al. (2013) Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency. Nat Cell Biol 15: 700-711

Wang X., Lu Z., Gomez A., Hon G. C., Yue Y., et al. (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505: 117-120

Wang Y. M., Zhou P., Wang L. Y., Li Z. H., Zhang Y. N., et al. (2012) Correlation between DNase I hypersensitive site distribution and gene expression in HeLa S3 cells. PLoS One 7: e42414

Warden C. D., Lee H., Tompkins J. D., Li X., Wang C., et al. (2013) COHCAP: an integrative genomic pipeline for single-nucleotide resolution DNA methylation analysis. Nucleic Acids Res 41: e117

Warnault V., Darcq E., Levine A., Barak S. and Ron D. (2013) Chromatin remodeling--a novel strategy to control excessive alcohol drinking. Transl Psychiatry 3: e231

Waszak S. M., Kilpinen H., Gschwind A. R., Orioli A., Raghav S. K., et al. (2014) Identification and removal of low-complexity sites in allele-specific analysis of ChIP-seq data. Bioinformatics 30: 165-171

Weber M., Davies J. J., Wittig D., Oakeley E. J., Haase M., et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37: 853-862

Wei Y., Chen R., Dimicoli S., Bueso-Ramos C., Neuberg D., et al. (2013) Global H3K4me3 genome mapping reveals alterations of innate immunity signaling and overexpression of JMJD3 in human myelodysplastic syndrome CD34+ cells. Leukemia 27: 2177-2186

Wei Z., Gao F., Kim S., Yang H., Lyu J., et al. (2013) Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. Cell Stem Cell 13: 36-47

Whisnant A. W., Bogerd H. P., Flores O., Ho P., Powers J. G., et al. (2013) In-depth analysis of the interaction of HIV-1 with cellular microRNA biogenesis and effector mechanisms. MBio 4: e000193

Will B., Vogler T. O., Bartholdy B., Garrett-Bakelman F., Mayer J., et al. (2013) Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. Nat Immunol 14: 437-445

Wolchinsky Z., Shivtiel S., Kouwenhoven E. N., Putin D., Sprecher E., et al. (2014) Angiomodulin is required for cardiogenesis of embryonic stem cells and is maintained by a feedback loop network of p63 and Activin-A. Stem Cell Res 12: 49-59

Woyke T., Tighe D., Mavromatis K., Clum A., Copeland A., et al. (2010) One bacterial cell, one complete genome. PLoS ONE 5: e10314

Woyke T., Sczyrba A., Lee J., Rinke C., Tighe D., et al. (2011) Decontamination of MDA reagents for single cell whole genome amplification. PLoS ONE 6: e26161

Xi Y., Yao J., Chen R., Li W. and He X. (2011) Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res 21: 718-724

Xi Y., Bock C., Muller F., Sun D., Meissner A., et al. (2012) RRBSMAP: a fast, accurate and user-friendly alignment tool for reduced representation bisulfite sequencing. Bioinformatics 28: 430-432

Xie W., Schultz M. D., Lister R., Hou Z., Rajagopal N., et al. (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 153: 1134-1148

Xing J., Witherspoon D. J. and Jorde L. B. (2013) Mobile element biology: new possibilities with high-throughput sequencing. Trends Genet 29: 280-289

Xu D., Shen W., Guo R., Xue Y., Peng W., et al. (2013) Top3beta is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. Nat Neurosci 16: 1238-1247

Xue Z., Huang K., Cai C., Cai L., Jiang C. Y., et al. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500: 593-597

Yadav S. P., Hao H., Yang H. J., Kautzmann M. A., Brooks M., et al. (2013) The transcription-splicing protein NonO/p54nrb and three NonO-interacting proteins bind to distal enhancer region and augment rhodopsin expression. Hum Mol Genet

Yadon A. N. and Tsukiyama T. (2013) DNA looping-dependent targeting of a chromatin remodeling factor. Cell Cycle 12: 1809-1810

Yaffe E. and Tanay A. (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet 43: 1059-1065

Yamaguchi S., Hong K., Liu R., Shen L., Inoue A., et al. (2012) Tet1 controls meiosis by regulating meiotic gene expression. Nature 492: 443-447

Yamaguchi S., Hong K., Liu R., Inoue A., Shen L., et al. (2013) Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. Cell Res 23: 329-339

Yamaguchi S., Shen L., Liu Y., Sendler D. and Zhang Y. (2013) Role of Tet1 in erasure of genomic imprinting. Nature 504: 460-464

Yamane A., Robbiani D. F., Resch W., Bothmer A., Nakahashi H., et al. (2013) RPA Accumulation during Class Switch Recombination Represents 5'-3' DNA-End Resection during the S-G2/M Phase of the Cell Cycle. Cell Rep 3: 138-147

Yan L., Yang M., Guo H., Yang L., Wu J., et al. (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 20: 1131-1139

Yano M., Hayakawa-Yano Y., Mele A. and Darnell R. B. (2010) Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. Neuron 66: 848-858

Yeh H. H., Young D., Gelovani J. G., Robinson A., Davidson Y., et al. (2013) Histone deacetylase class II and acetylated core histone immunohistochemistry in human brains with Huntington's disease. Brain Res 1504: 16-24

Yilmaz S. and Singh A. K. (2011) Single cell genome sequencing. Curr Opin Biotechnol 23: 437-443

Yokobayashi S., Liang C. Y., Kohler H., Nestorov P., Liu Z., et al. (2013) PRC1 coordinates timing of sexual differentiation of female primordial germ cells. Nature 495: 236-240

Yu M., Hon G. C., Szulwach K. E., Song C. X., Jin P., et al. (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. Nat Protoc 7: 2159-2170

Yu M., Hon G. C., Szulwach K. E., Song C. X., Zhang L., et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the Mammalian genome. Cell 149: 1368-1380

Yu M., Hon G. C., Szulwach K. E., Song C. X., Zhang L., et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 149: 1368-1380

Yu W., Jin C., Lou X., Han X., Li L., et al. (2011) Global analysis of DNA methylation by Methyl-Capture sequencing reveals epigenetic control of cisplatin resistance in ovarian cancer cell. PLoS One 6: e29450

Zarnack K., Konig J., Tajnik M., Martincorena I., Eustermann S., et al. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell 152: 453-466

Zentner G. E. and Henikoff S. (2012) Surveying the epigenomic landscape, one base at a time. Genome Biol 13: 250

Zhan Q., Fang Y., He Y., Liu H. X., Fang J., et al. (2012) Function annotation of hepatic retinoid x receptor alpha based on genome-wide DNA binding and transcriptome profiling. PLoS One 7: e50013

Zhang B., Zhou Y., Lin N., Lowdon R. F., Hong C., et al. (2013) Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Res 23: 1522-1540

Zhang C., Frias M. A., Mele A., Ruggiu M., Eom T., et al. (2010) Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. Science 329: 439-443

Zhang C. and Darnell R. B. (2011) Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol 29: 607-614

Zhang J., Poh H. M., Peh S. Q., Sia Y. Y., Li G., et al. (2012) ChIA-PET analysis of transcriptional chromatin interactions. Methods 58: 289-299

Zhang W., Wu Y., Schnable J. C., Zeng Z., Freeling M., et al. (2012) High-resolution mapping of open chromatin in the rice genome. Genome Res 22: 151-162

Zhang W., Zhang T., Wu Y. and Jiang J. (2012) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. Plant Cell 24: 2719-2731

Zhang X., Robertson G., Woo S., Hoffman B. G. and Gottardo R. (2012) Probabilistic Inference for Nucleosome Positioning with MNase-Based or Sonicated Short-Read Data. PLoS ONE 7: e32095

Zhang Y., McCord R. P., Ho Y. J., Lajoie B. R., Hildebrand D. G., et al. (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 148: 908-921

Zhao J., Ohsumi T. K., Kung J. T., Ogawa Y., Grau D. J., et al. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell 40: 939-953

Zhao Y., Guo S., Sun J., Huang Z., Zhu T., et al. (2012) Methylcap-seq reveals novel DNA methylation markers for the diagnosis and recurrence prediction of bladder cancer in a Chinese population. PLoS ONE 7: e35175

Zhao Z., Tavoosidana G., Sjolinder M., Gondor A., Mariano P., et al. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38: 1341-1347

Zilbauer M., Rayner T. F., Clark C., Coffey A. J., Joyce C. J., et al. (2013) Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type-specific hypomethylated regions. Blood 122: e52-60

Ziller M. J., Muller F., Liao J., Zhang Y., Gu H., et al. (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genet 7: e1002389

Zomer A., Burghout P., Bootsma H. J., Hermans P. W. and van Hijum S. A. (2012) ESSENTIALS: Software for Rapid Analysis of High Throughput Transposon Insertion Sequencing Data. PLoS ONE 7: e43012

Zong C., Lu S., Chapman A. R. and Xie X. S. (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338: 1622-1626

Zund D., Gruber A. R., Zavolan M. and Muhlemann O. (2013) Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat Struct Mol Biol 20: 936-943

MasterPure™ Complete DNA and RNA Purification Kit

epicentre®
an Illumina® company

Pure Nucleic Acids

MasterPure™ Complete DNA and RNA Purification kit

- Extract and purify total nucleic acids (TNA), DNA or RNA
- Pure for sequencing, qPCR and other molecular biology applications
- Scalable reactions
- High purity and yield
- Non-Toxic

The MasterPure™ Complete Kit purifies high yields of intact total nucleic acid, DNA, or RNA. MasterPure is suitable for every type of biological material.

MasterPure is optimized for use with:

- Illumina® sequencing
- qPCR
- PCR
- Molecular biology applications

Table 1. Purify any sample.

Sample	Sample Size	TNA µg	DNA µg	RNA µg
HeLa/HL60 cells	1 X 10^6 cells	10-30	3-12	7-15
Liver	5 mg	33-42	5-10	13-25
Brain	5 mg	9-13	6-9	4-11
Heart	5 mg	6-10	4-7	4-5
Blood	200 µl	3-10	3-9	
Buffy coat	300 µl	40-55	40-55	3-6
E. coli	3.5 x 10^6 cells	2.5-2.8	1.3-1.6	1.6-1.8
Yeast*	2.2 x 10^6 cells			11-18
(*S. cerevisiae*)	1.1 x 10^7 cells			70-78

Many different, diverse sample types have been purified by MasterPure. Several are shown in Table 1. MasterPure is available for virtually any type of sample.

MasterPure may be used to purify total nucleic acid, DNA or RNA from any sample. Total nucleic acid purification permits you to compare DNA and RNA from the same sample to gain a deeper understanding of your sample.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089

PDS 010

Figure 2. MasterPure easily purifies many different sample types.

MasterPure has been published showing excellent purification of nearly every type of sample. Here we see just a few of the types of samples that have been published using MasterPure.

Purify any sample

MasterPure has been shown to work for many types of human tissue and blood samples, plants, and bacteria.

MasterPure is safe and nontoxic. No dangerous chemicals, phenol or hazards are used in the method. MasterPure is a wise choice for safety and high yields of RNA, DNA, or total nucleic acids.

Stop stocking three different kits for small, moderate and abundant samples! MasterPure is designed to be used with small, moderate and abundant samples without the need for many kits. One MasterPure kit permits you to purify RNA, DNA or total nucleic acid from any amount of sample.

One kit to purify your choice of nucleic acids.

Suitable for Illumina® sequencing

Total nucleic acid, DNA or RNA purified by MasterPure is suitable for use with Illumina sequencing. All sequencing applications begin with MasterPure, including:

- Ribo-Zero
- RNA-Seq
- Bisulfite sequencing for epigenetics
- DNA-Seq
- Exome capture
- More...

Cat. #	Quantity
MasterPure™ Complete DNA and RNA Purification kit	
MC85200	200 DNA Purifications 100 RNA Purifications
MC89010	10 DNA Purifications 5 RNA Purifications
MasterPure™ DNA Purification Kit	
MCD85201	200 Purifications

DNA-Sequencing	
Description	**Catalog Number**
MasterPure™ Complete DNA and RNA Purification Kit	MC85200
MasterPure™ DNA Purification Kit	MCD85201
TruSeq DNA PCR-Free LT Sample Preparation Kit - Set A	FC-121-3001
TruSeq DNA PCR-Free LT Sample Preparation Kit - Set B	FC-121-3002
TruSeq DNA PCR-Free HT Sample Preparation Kit	FC-121-3003
TruSeq Nano DNA LT Sample Preparation Kit - Set A	FC-121-4001
TruSeq Nano DNA LT Sample Preparation Kit - Set B	FC-121-4002
TruSeq Nano DNA HT Sample Preparation Kit	FC-121-4003
Nextera Rapid Capture Exome (8 rxn x 1 Plex)	FC-140-1000
Nextera Rapid Capture Exome (8 rxn x 3 Plex)	FC-140-1083
Nextera Rapid Capture Exome (8 rxn x 6 Plex)	FC-140-1086
Nextera Rapid Capture Exome (8 rxn x 9 Plex)	FC-140-1089
Nextera Rapid Capture Exome (2 rxn x 12 Plex)	FC-140-1001
Nextera Rapid Capture Exome (4 rxn x 12 Plex)	FC-140-1002
Nextera Rapid Capture Exome (8 rxn x 12 Plex)	FC-140-1003
Nextera Rapid Capture Expanded Exome (2 rxn x 12 Plex)	FC-140-1004
Nextera Rapid Capture Expanded Exome (4 rxn x 12 Plex)	FC-140-1005
Nextera Rapid Capture Expanded Exome (8 rxn x 12 Plex)	FC-140-1006
EpiGnome™ Methyl-Seq Kit	EGMK81312

ChIP	
Description	**Catalog Number**
TruSeq ChIP Sample Preparation Kit - Set A	IP-202-1012
TruSeq ChIP Sample Preparation Kit - Set B	IP-202-1024

Methylation Arrays	
Description	**Catalog Number**
HumanMethylation450 DNA Analysis BeadChip Kit (24 samples)	WG-314-1003
HumanMethylation450 DNA Analysis BeadChip Kit (48 samples)	WG-314-1001
HumanMethylation450 DNA Analysis BeadChip Kit (96 samples)	WG-314-1002

Data Sheet: Illumina® Sequencing

illumina®

TruSeq™ RNA and DNA Sample Preparation Kits v2

Master-mixed reagents, optimized adapter design, and a flexible workflow provide a simple, cost-effective method for preparing RNA and DNA samples for scalable next-generation sequencing.

Highlights

- **Simple Workflow for RNA and DNA:** Master-mixed reagents and minimal hands-on steps.
- **Scalable and Cost-Effective Solution:** Optimized formulations and plate-based processing enables large-scale studies at a lower cost.
- **Enhanced Multiplex Performance:** Twenty-four adaptor-embedded indexes enable high-throughput processing and greater application flexibility.
- **High-Throughput Gene Expression Studies:** Gel-free, automation-friendly RNA sample preparation for rapid expression profiling.

Introduction

Illumina next-generation sequencing (NGS) technologies continue to evolve, offering increasingly higher output in less time. Keeping pace with these developments requires improvements in sample preparation. To maximize the benefits of NGS and enable delivery of the highest data accuracy, Illumina offers the TruSeq RNA and DNA Sample Preparation Kits (Figure 1).

The TruSeq RNA and DNA Sample Preparation Kits provide a simple, cost-effective solution for generating libraries from total RNA or genomic DNA that are compatible with Illumina's unparalleled sequencing output. Master-mixed reagents eliminate the majority of pipetting steps and reduce the amount of clean-up, as compared to previous methods, minimizing hands-on time. New automation-friendly workflow formats enable parallel processing of up to 96 samples. This results in economical, high-throughput RNA or DNA sequencing studies achieved with the easiest-to-use sample preparation workflow offered by any NGS platform.

Simple and Cost-Effective Solution

Whether processing samples for RNA-Seq, genomic sequencing, or exome enrichment, the TruSeq kits provide significantly improved library preparation over previously used methods. New protocols reduce the number of purification, sample transfer, and pipetting steps. The new universal, methylated adaptor design incorporates an index sequence at the initial ligation step for improved workflow efficiency and more robust multiplex sequencing. For maximum flexibility, the same TruSeq kit can be used to prepare samples for single-read, paired-end, and multiplexed sequencing on all Illumina sequencing instruments.

TruSeq DNA and RNA Sample Prep kits include gel-free protocols that eliminate the time-intensive gel purification step found in other methods, making the process more consistent and fully automatable.

The gel-free protocol for TruSeq DNA sample preparation is available for target enrichment using the TruSeq Exome Enrichment or TruSeq Custom Enrichment kits.

TruSeq sample preparation makes RNA sequencing for high-throughput experiments more affordable, enabling gene expression profiling studies to be performed with NGS at a lower cost than arrays. It also provides a cost-effective DNA sequencing solution for large-scale whole-genome resequencing, targeted resequencing, *de novo* sequencing, metagenomics, and methlyation studies.

Enhanced Multiplex Performance

TruSeq kits take advantage of improved multiplexing capabilities to increase throughput and consistency, without compromising results. Both the RNA and DNA preparation kits include adapters containing unique index sequences that are ligated to sample fragments at the beginning of the library construction process. This allows the samples to be pooled and then individually identified during downstream analysis. The result is a more efficient, streamlined workflow that leads directly into a superior multiplexing solution. There are no additional PCR steps required for index incorporation, enabling a robust, easy-to-follow procedure. With 24 unique indexes available, up to 384 samples can be processed in parallel on a single HiSeq 2000 run.

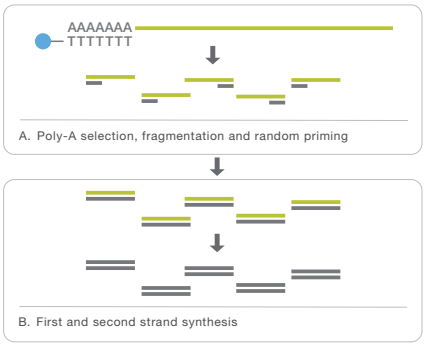
TruSeq RNA Sample Preparation

With TruSeq reagents, researchers can quickly and easily prepare samples for next-generation sequencing (Figure 2). Improvements in the RNA to cDNA conversion steps have significantly enhanced the overall workflow and performance of the assay (Figure 3).

Figure 1: TruSeq Sample Preparation Kits

TruSeq Sample Preparation Kits are available for both genomic DNA and RNA samples.

Figure 2: Optimized TruSeq RNA Sample Preparation



Starting with total RNA, mRNA is polyA-selected and fragmented. It then undergoes first- and second-strand synthesis to produce products ready for library construction (Figure 4).

Starting with total RNA, the messenger RNA is first purified using polyA selection (Figure 2A), then chemically fragmented and converted into single-stranded cDNA using random hexamer priming. Next, the second strand is generated to create double-stranded cDNA (Figure 2B) that is ready for the TruSeq library construction workflow (Figure 4).

Efficiencies gained in the polyA selection process, including reduced sample transfers, removal of precipitation steps, and combining of elution and fragmentation into a single step, enable parallel processing of up to 48 samples in approximately one hour. This represents a 75% reduction in hands-on time for this portion of library construction. Improving performance, the optimized random hexamer priming strategy provides the most even coverage across transcripts, while allowing user-defined adjustments for longer or shorter insert lengths.

Eliminating all column purification and gel selection steps from the workflow removes the most time-intensive portions, while improving the assay robustness. It also allows for decreased input levels of RNA—as low as 100 ng— and maintains single copy per gene sensitivity.

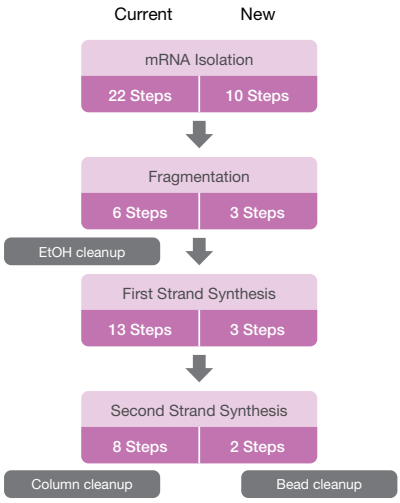
TruSeq DNA Sample Preparation

The TruSeq DNA Sample Preparation Kits are used to prepare DNA libraries with insert sizes from 300–500 bp for single, paired-end, and multiplexed sequencing. The protocol supports shearing by either sonication or nebulization with a low input requirement of 1 ug of DNA.

Sequence-Ready Libraries

Library construction begins with either double-stranded cDNA synthesized from RNA or fragmented gDNA (Figure 4A). Blunt-end DNA fragments are generated using a combination of fill-in reactions and exonuclease activity (Figure 4B). An 'A'- base is then added to the blunt ends of each strand, preparing them for ligation to the sequencing adapters (Figures 4C). Each adapter contains a 'T'-base overhang on 3'-end providing a complementary overhang for ligating the adapter

Figure 3: TruSeq RNA Sample Preparation Reagents Provide Significant Savings in Time and Effort



	Current Methods	TruSeq Methods	Savings
No. of Steps	49	18	31
Time (hours)	16	12	25%

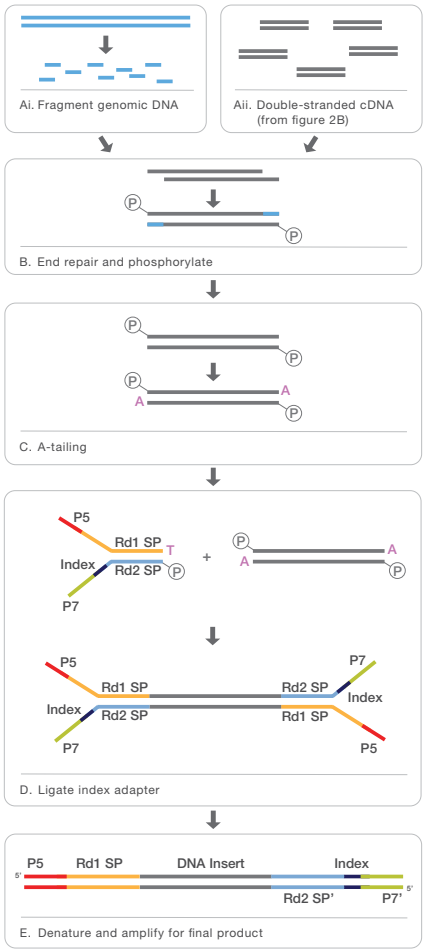
Compared to current methods for preparing mRNA samples for sequencing, use of the TruSeq reagents significantly reduces the number of steps and hands-on time.

Table 1: Savings When Processing 96 Samples

> 50% of pipetting steps eliminated
> 50% of reagent tubes eliminated
> 75% of clean-up steps eliminated
> 50% of sample transfer steps eliminated

Compared to previous kits, processing multiple samples with the new TruSeq Sample Preparation Kits provides significant reductions in library construction costs, the number of steps, hands-on time, and PCR dependency.

Figure 4: Adapter Ligation Results in Sequence-Ready Constructs without PCR



Library construction begins with either fragmented genomic DNA or double-stranded cDNA produced from total RNA (Figure 4A). Blunt-end fragments are created (Figure 4B) and an A-base is then added (Figure 4C) to prepare for indexed adapter ligation (Figure 4D). Final product is created (Figure 4E), which is ready for amplification on either the cBot or the Cluster Station.

to the A-tailed fragmented DNA. These newly redesigned adapters contain the full complement of sequencing primer hybridization sites for single, paired-end, and multiplexed reads. This eliminates the need for additional PCR steps to add the index tag and multiplex primer sites (Figure 4D). Following the denaturation and amplification steps (Figure 4E), libraries can be pooled with up to 12 samples per lane (96 sample per flow cell) for cluster generation on either cBot or the Cluster Station.

Master-mixed reagents and an optimized protocol improve the library construction workflow, significantly decreasing hands-on time and reducing the number of clean-up steps when processing samples for large-scale studies (Table 1). The simple and scalable workflow allows for high-throughput and automation-friendly solutions, as well as simultaneous manual processing for up to 96 samples. In addition, enhanced troubleshooting features are incorporated into each step of the workflow, with quality control sequences supported by Illumina RTA software.

Enhanced Quality Controls

Specific Quality Control (QC) sequences, consisting of double-stranded DNA fragments, are present in each enzymatic reaction of the TruSeq sample preparation protocol: end repair, A-tailing, and ligation. During analysis, the QC sequences are recognized by the RTA software (versions 1.8 and later) and isolated from the sample data. The presence of these controls indicates that its corresponding step was successful. If a step was unsuccessful, the control sequences will be substantially reduced. QC controls assist in comparison between experiments and greatly facilitate troubleshooting.

Designed For Automation

The TruSeq Sample Preparation Kits are compatible with high-throughput, automated processing workflows. Sample preparation can be performed in standard 96-well microplates with master-mixed reagent pipetting volumes optimized for liquid-handling robots. Barcodes on reagents and plates allow end-to-end sample tracking and ensure that the correct reagents are used for the correct protocol, mitigating potential tracking errors.

Part of an Integrated Sequencing Solution

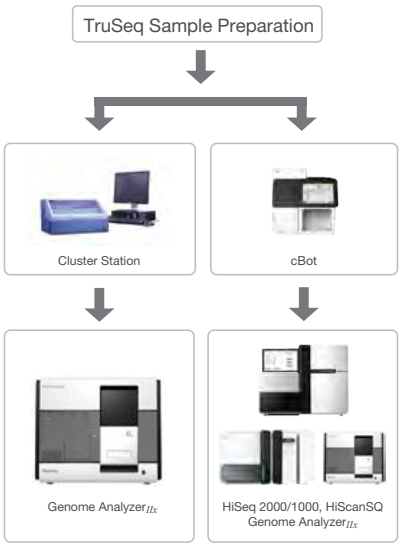
Samples processed with the TruSeq Sample Preparation Kits can be amplified on either the cBot Automated Cluster Generation System or the Cluster Station and used with any of Illumina's next-generation sequencing instruments, including HiSeq™ 2000, HiSeq 1000, HiScan™SQ, Genome Analyzer$_{IIx}$ (Figure 5).

Summary

Illumina's new TruSeq Sample Preparation Kits enable simplicity, convenience, and affordability for library preparation. Enhanced multiplexing with 24 unique indexes allows efficient high-throughput processing. The pre-configured reagents, streamlined workflow, and automation-friendly protocol save researchers time and effort in their next-generation sequencing pursuits, ultimately leading to faster discovery and publication.

Learn more about Illumina's next-generation sequencing solutions at www.illumina.com/sequencing.

Figure 5. Illumina's Complete Sequencing Solution



The TruSeq Sample Preparation Kits readily fit in with Illumina's advanced next-generation sequencing solutions.

Ordering Information

Product	Catalog No.
For RNA Preparation	
TruSeq RNA Sample Preparation Kit v2, Set A (12 indexes, 48 samples)	RS-122-2001
TruSeq RNA Sample Preparation Kit v2, Set B (12 indexes, 48 samples)	RS-122-2002
For DNA Preparation	
TruSeq DNA Sample Preparation Kit v2, Set A (12 indexes, 48 samples)	FC-121-2001
TruSeq DNA Sample Preparation Kit v2, Set B (12 indexes, 48 samples)	FC-121-2002
For Cluster Generation on cBot and Sequencing on the HiSeq 2000/1000 and HiScanSQ	
TruSeq Paired-End Cluster Kit v3—cBot—HS (1 flow cell)	PE-401-3001
TruSeq Single-Read Cluster Kit v3—cBot—HS (1 flow cell)	GD-401-3001
For Cluster Generation on cBot and Sequencing on the Genome Analyzer$_{IIx}$	
TruSeq Paired-End Cluster Kit v2—cBot—GA (1 flow cell)	PE-300-2001
TruSeq Single-Read Cluster Kit v2—cBot—GA (1 flow cell)	GD-300-2001
For Cluster Generation on the Cluster Station and Sequencing on the Genome Analyzer$_{IIx}$	
TruSeq Paired-End Cluster Kit v5—CS—GA (1 flow cell)	PE-203-5001
TruSeq Single-Read Cluster Kit v5—CS—GA (1 flow cell)	GD-203-5001

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

Pub. No. 970-2009-039 Current as of 27 April 2011

Data Sheet: Sequencing

illumina®

TruSeq® DNA PCR-Free Sample Preparation Kit

Setting new standards for unbiased data quality and superior coverage.

Highlights

- **Superior Coverage**
  Elimination of PCR-induced bias and fewer coverage gaps provide greater access to the genome
- **Faster Sample Preparation**
  PCR-Free protocol accelerates the most widely adopted sample preparation chemistry
- **Unprecedented Flexibility**
  PCR-Free kits are optimized to support a variety of read lengths and applications
- **Inclusive Solution**
  Reliable solution includes master-mixed reagents, size-selection beads, and up to 96 indices for the highest operational efficiency

Introduction

The TruSeq DNA PCR-Free Sample Preparation Kit offers numerous enhancements to the industry's most widely adopted sample preparation workflow, providing an optimized, all-inclusive sample preparation for whole-genome sequencing applications. By eliminating PCR amplification steps, the PCR-Free protocol removes typical PCR-induced bias and streamlines the proven TruSeq workflow. This results in excellent data quality and detailed sequence information for traditionally challenging regions of the genome. Two kit types are available to accommodate a range of study designs: the TruSeq DNA PCR-Free LT Sample Preparation Kit for low-throughput studies and the TruSeq DNA PCR-Free HT Sample Preparation Kit for high-throughput studies (Figure 1).

Accelerated Sample Preparation

The TruSeq DNA sample preparation workflow has been streamlined further by removing the PCR step and replacing gel-based size selection with bead-based selection (Figure 2). This kit offers unprecendented flexibility with two protocol options for generating either large (550 bp) or small (350 bp) insert sizes to support a variety of applications, matching the ever-increasing read lengths of Illumina sequencing instruments. Master-mixed reagents, provided sample purification beads, and optimized protocols contribute to the simplified library construction workflow, requiring minimal hands-on time and few cleanup steps for processing large sample numbers. TruSeq DNA PCR-Free sample preparation decreases library preparation time, empowering applications from microbial sequencing to whole human genome sequencing.[1]

Figure 1: TruSeq DNA PCR-Free Sample Preparation Kit

TruSeq DNA PCR-Free kits are an efficient solution for preparing and indexing sample libraries. The TruSeq DNA PCR-Free LT kit provides up to 24 indices for low-throughput studies (with both Sets A and B), while the TruSeq DNA PCR-Free HT kit includes 96 dual-index combinations for high-throughput studies.

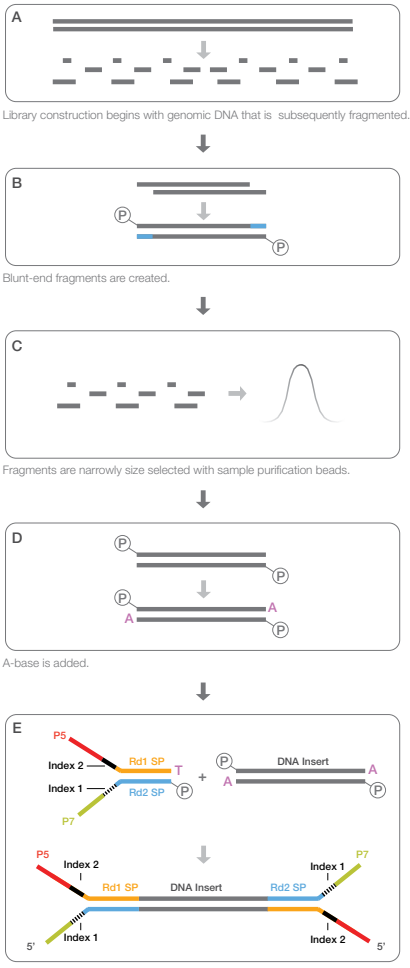
Innovative Sample Preparation Chemistry

TruSeq DNA PCR-Free Sample Preparation kits are used to prepare DNA libraries for single, paired-end, and indexed sequencing. The protocol supports shearing by Covaris ultrasonication, requiring 1 µg of input DNA for an average insert size of 350 bp or 2 µg for an average insert size of 550 bp. Library construction begins wtih fragmented gDNA (Figure 2A). Blunt-end DNA fragments are generated using a combination of fill-in reactions and exonuclease activity (Figure 2B), and size selection is performed with provided sample purification beads (Figure 2C). An A-base is then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters (Figure 2D). Each adapter contains a T-base overhang for ligating the adapter to the A-tailed fragmented DNA. These adapters contain the full complement of sequencing primer hybridization sites for single, paired-end, and indexed reads. With no need for additional PCR amplification, single or dual-index adapters are ligated to the fragments and samples are ready for cluster generation (Figure 2E).

Superior Coverage

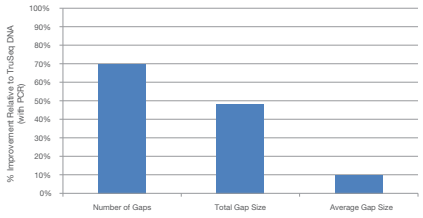
The TruSeq DNA PCR-Free Sample Preparation Kit optimizes sequencing data to provide greater insight into the genome, including coding, regulatory, and intronic regions. PCR-Free sample preparation generates reduced library bias and gaps (Figure 3). Exceptional data quality delivers base-pair resolution of somatic and *de novo* mutations, supporting accurate identification of causative variants. The removal of PCR amplification from the TruSeq workflow removes amplification biases to improve coverage uniformity across the genome (Figure 4).

Figure 2: Adapter Ligation Results in Sequence-Ready Constructs without PCR



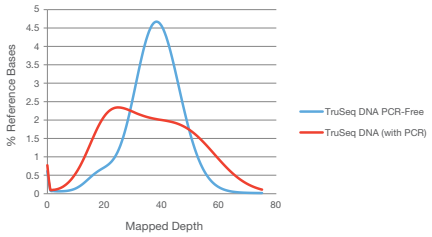
*The TruSeq DNA PCR-Free LT indexing solution features a single-index adapter at this step.

Figure 3: Fewer Gaps in Coverage



TruSeq DNA PCR-Free libraries show significant reduction in the number and total size of gaps when compared to libraries prepared using the TruSeq DNA (with PCR) protocol. A gap is defined as a region ≥ 10 bp in length, where an accurate genotype cannot be determined due to low depth, low alignment scores, or low base quality.

Figure 4: Greater Coverage Uniformity



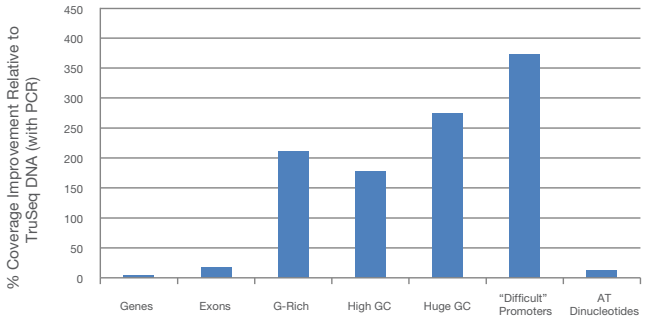
TruSeq DNA PCR-Free libraries provide greater coverage uniformity across the genome when compared to those generated using the TruSeq DNA protocol.

The PCR-Free kit also provides superior coverage of traditionally challenging genomic content, including GC-rich regions, promoters, and repetitive regions (Figure 5), allowing researchers to access more genomic information from each sequencing run (Figure 6).

Efficient Sample Multiplexing

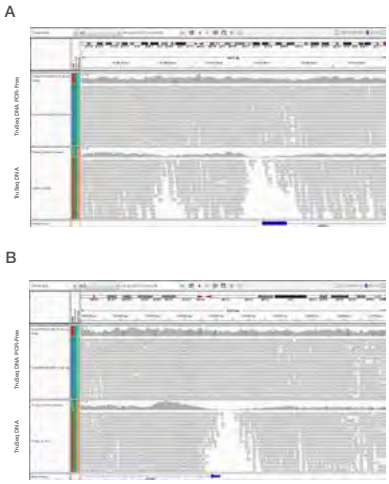
TruSeq DNA PCR-Free Sample Preparation kits provide an innovative solution for sample multiplexing. Indices are added to sample gDNA fragments using a simple PCR-Free procedure. For the greatest operational efficiency, up to 96 pre-plated, uniquely indexed samples can be pooled and sequenced together in a single flow cell lane on any Illumina sequencing platform. After sequencing, the indices are used to demultiplex the data and accurately assign reads to the proper sample in the pool. The TruSeq DNA PCR-Free LT kit uses a single index for demultiplexing, while the TruSeq DNA PCR-Free HT kit employs a dual-indexing strategy, using a unique combination of two indices to demultiplex.

Figure 5: Increased Coverage of Challenging Regions



When compared to libraries generated by PCR-based workflows, such as TruSeq DNA Sample Preparation, PCR-Free libraries show improved coverage for challenging regions of the genome. These regions include known human protein coding and non-protein coding exons and genes defined in the RefSeq Genes track in the UCSC Genome Browser.[2] G-Rich regions denote 30 bases with ≥ 80% G. High GC regions are defined as 100 bases with ≥ 75% GC content. Huge GC regions are defined as 100 bases with ≥ 85% GC content. "Difficult" promoters denote the set of 100 promoter regions that are insufficiently covered, which have been empirically defined by the Broad Institute of MIT and Harvard.[3] AT dinucleotides indicate 30 bases of repeated AT dinucleotide.

Figure 6: PCR-Free Protocol Eliminates Coverage Gaps in GC-Rich Content



Increased coverage of TruSeq DNA PCR-Free libraries results in fewer coverage gaps, demonstrated here in the GC-rich coding regions of the *RNPEPL1* promoter (A) and the *CREBBP* promoter (B). PCR-Free sequence information is shown in the top panels of A and B, while sequence data generated using TruSeq DNA protocol (with PCR) are shown in the lower panels.

The TruSeq LT kit includes up to 24 indices with two sets of 12each, and the TruSeq HT kit offers 96 indices for efficient experimental design.

Multi-sample studies can be conveniently managed using the Illumina Experiment Manager, a freely available software tool that provides easy reaction setup for plate-based processing. It allows researchers to quickly configure the index sample sheet (i.e., sample multiplexing matrix) for the instrument run, enabling automatic demultiplexing.

Flexible and Inclusive Sample Preparation

The TruSeq family of sample preparation solutions offers several kits for sequencing applications, compatible with a range of research needs and study designs (Table 1). All TruSeq kits support high- and low-throughput studies. The TruSeq DNA PCR-Free kit provides superior coverage quality and drastically reduces library bias and coverage gaps, without requiring PCR amplification. These kits enhance the industry's most widely adopted DNA sample preparation method, empowering next-generation sequencing applications.

Simplified Solution

The comprehensive solution includes sample preparation reagents, sample purification beads, and robust TruSeq barcodes for sample multiplexing, providing a complete preparation method optimized for the highest performance on all Illumina sequencing platforms. The TruSeq DNA PCR-Free kit leverages the flexibility of two kit options, 24-sample and 96-sample, for a scalable experimental approach. With a simplified workflow and multiplexing options, the TruSeq DNA PCR-Free protocol offers the fastest library preparation method for the highest data quality.

Table 1: TruSeq DNA Sample Preparation Kits

Specification	TruSeq Nano DNA	TruSeq DNA PCR-Free	TruSeq DNA
Description	Based upon widely adopted TruSeq sample prep, with lower input and improved data quality	Superior genomic coverage with radically reduced library bias and gaps	Original TruSeq next-generation sequencing sample preparation method
Input quantity	100–200 ng	1–2 µg	1 µg
Includes PCR	Yes	No	Yes
Assay time	~6 hours	~5 hours	1–2 days
Hands-on time	~5 hours	~4 hours	~8 hours
Target insert size	350 bp or 550 bp	350 bp or 550 bp	300 bp
Gel-Free	Yes	Yes	No
Number of samples supported	24 (LT) or 96 (HT) samples	24 (LT) or 96 (HT) samples	48 (LT) or 96 (HT) samples
Supports enrichment	No*	No*	Yes
Size-selection beads	Included	Included	Not included
Applications	Whole-genome sequencing applications, including whole-genome resequencing, *de novo* assembly, and metagenomics studies		
Sample multiplexing	24 single indices or 96 dual-index combinations		
Compatible Illumina sequencers	HiSeq®, HiScanSQ™, Genome Analyzer™, and MiSeq® systems		

*Nextera Rapid Capture products support a variety of enrichment applications. For more information, visit www.illumina.com/NRC.

Summary

The TruSeq DNA PCR-Free Sample Preparation Kit optimizes the TruSeq workflow to deliver a faster sample preparation method for any species. The choice between protocol options provides greater flexibility to support a variety of applications and genomic studies. The PCR-Free kit also removes PCR-induced bias to facilitate detailed and accurate insight into the genome. By leveraging a faster workflow and superior data quality, the TruSeq DNA PCR-Free Sample Preparation Kit enables researchers to obtain high-quality genomic data, faster.

References

1. Saunders CJ, Miller NA, Soden SE, Dinwiddie DL, Noll A, et al. (2012) Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. Science Translational Medicine 4(154): 154ra135.
2. genome.ucsc.edu
3. www.broadinstitute.org

Ordering Information

Product	Catalog No.
TruSeq DNA PCR-Free LT Sample Preparation Kit Set A (24 samples)	FC-121-3001
TruSeq DNA PCR-Free LT Sample Preparation Kit Set B (24 samples)	FC-121-3002
TruSeq DNA PCR-Free HT Sample Preparation Kit (96 samples)	FC-121-3003

Illumina • 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. 770-2013-001 Current as of 16 May 2013

illumina®

Data Sheet: Sequencing

illumina®

TruSeq® Nano DNA Sample Preparation Kit

A low-input method that delivers a high-confidence, comprehensive view of the genome for virtually any sequencing application.

Highlights

- **Low Sample Input**
  Excellent data quality from as little as 100 ng input empowers interrogation of samples with limited available DNA
- **Excellent Coverage Quality**
  Significantly reduced library bias and gaps in coverage provide greater insight into the genome
- **Unprecedented Flexibility**
  Streamlined TruSeq workflow enables library preparation in less than one day, while supporting a variety of read lengths
- **Inclusive Solution**
  Reliable solution includes master-mixed reagents, size-selection beads, and up to 96 indices for the highest operational efficiency

Figure 1: TruSeq Nano DNA Sample Preparation Kit

TruSeq Nano DNA Sample Preparation Kits offer a low-input solution for preparing and indexing sample libraries. The TruSeq Nano DNA LT kit provides up to 24 indices for low-throughput studies (with both Sets A and B), while the TruSeq Nano DNA HT kit includes 96 dual-index combinations for high-throughput studies.

Introduction

By offering a low-input method based on the industry's most widely adopted sample preparation workflow, the TruSeq Nano DNA Sample Preparation Kit enables efficient interrogation of samples that have limited available DNA. This kit significantly reduces typical PCR-induced bias and provides detailed sequence information for traditionally challenging regions of the genome. Two kit types are available to accommodate a range of study designs: the TruSeq Nano DNA LT Sample Preparation Kit for low-throughput studies and the TruSeq Nano DNA HT Sample Preparation Kit for high-throughput studies (Figure 1).

Low Sample Input

The TruSeq Nano DNA protocol eliminates the typical requirement for micrograms of DNA, enabling researchers to study samples with limited available DNA (e.g., tumor samples) and supporting preservation of samples for use in future or alternate studies. This kit offers the flexibility of two protocols for generating large (550 bp) or small (350 bp) insert sizes to support a diverse range of applications. In addition to accelerating the workflow, simple bead-based size selection avoids typical sample loss associated with gel-based selection. TruSeq Nano DNA kits are validated for high-quality genomic coverage for virtually any whole-genome sequencing application.

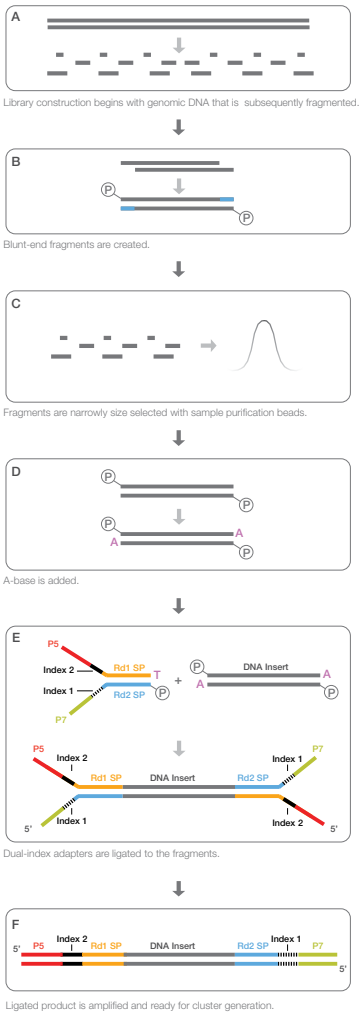
Accelerated Sample Preparation

The TruSeq DNA sample preparation workflow has been streamlined by replacing gel-based size selection with bead-based selection (Figure 2), enabling researchers to prepare high-quality libraries in less than a day. Optimized for a variety of read lengths, from 2 × 101 bp to 2 × 151 bp, the TruSeq Nano DNA kit is designed to match the ever-increasing read lengths of Illumina sequencing instruments. Master-mixed reagents, provided sample purification beads for cleanup and size selection, robust TruSeq indices, and optimized protocols contribute to the simplified workflow, requiring minimal hands-on time and few cleanup steps for processing large sample numbers.

Innovative Sample Preparation Chemistry

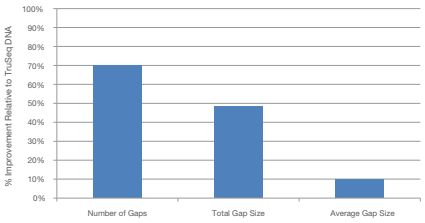
These kits are used to prepare DNA libraries for single-read, paired-end, and indexed sequencing. The TruSeq Nano DNA protocol supports shearing by Covaris ultrasonication, requiring 100 ng of input DNA for an average insert size of 350 bp or 200 ng DNA for an average insert size of 550 bp. Library construction begins with fragmented gDNA (Figure 2A). Blunt-end DNA fragments are generated using a combination of fill-in reactions and exonuclease activity (Figure 2B), and size selection is performed with provided sample purification beads (Figure 2C). An A-base is then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters (Figure 2D). Each adapter contains a T-base overhang for ligating the adapter to the A-tailed fragmented DNA. These adapters contain the full complement of sequencing primer hybridization sites for single, paired-end, and indexed reads. Single- or dual-index adapters are ligated to the fragments (Figure 2E) and the ligated products are amplified with reduced-bias PCR (Figure 2F).

Figure 2: TruSeq Nano DNA Workflow



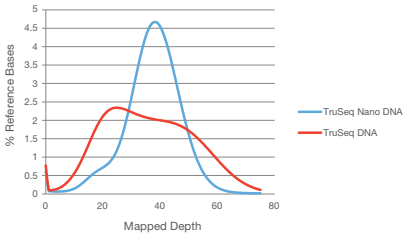
The TruSeq Nano DNA LT indexing solution features a single-index adapter at Step E.

Figure 3: Fewer Gaps in Coverage



TruSeq Nano DNA libraries show significant reduction in the number and total size of gaps when compared to libraries prepared using the TruSeq DNA protocol. A gap is defined as a region ≥ 10 bp in length, where an accurate genotype cannot be determined due to low depth, low alignment scores, or low base quality.

Figure 4: Greater Coverage Uniformity



TruSeq Nano DNA libraries provide greater coverage uniformity across the genome when compared to those generated using the TruSeq DNA protocol.

Table 1: TruSeq DNA Sample Preparation Kits

Specification	TruSeq Nano DNA	TruSeq DNA PCR-Free	TruSeq DNA
Description	Based upon widely adopted TruSeq sample prep, with lower input and improved data quality	Superior genomic coverage with radically reduced library bias and gaps	Original TruSeq next-generation sequencing sample preparation method
Input quantity	100–200 ng	1–2 µg	1 µg
Includes PCR	Yes	No	Yes
Assay time	~6 hours	~5 hours	1–2 days
Hands-on time	~5 hours	~4 hours	~8 hours
Target insert size	350 bp or 550 bp	350 bp or 550 bp	300 bp
Gel-Free	Yes	Yes	No
Number of samples supported	24 (LT) or 96 (HT) samples	24 (LT) or 96 (HT) samples	48 (LT) or 96 (HT) samples
Supports enrichment	No*	No*	Yes
Size-selection beads	Included	Included	Not included
Applications	Whole-genome sequencing applications, including whole-genome resequencing, *de novo* assembly, and metagenomics studies		
Sample multiplexing	24 single indices or 96 dual-index combinations		
Compatible Illumina sequencers	HiSeq®, HiScanSQ™, Genome Analyzer™, and MiSeq® systems		

*Nextera Rapid Capture products support a variety of enrichment applications. For more information, visit www.illumina.com/NRC.

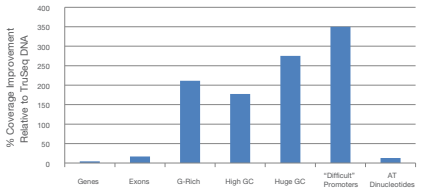
Excellent Coverage Quality

TruSeq Nano DNA kits reduce the number and average size of typical PCR-induced gaps in coverage (Figure 3), delivering exceptional data quality. The enhanced workflow reduces library bias and improves coverage uniformity across the genome (Figure 4). These kits also provide excellent coverage of traditionally challenging genomic content, including GC-rich regions, promoters, and repetitive regions (Figure 5). High data quality delivers base-pair resolution, providing a detailed view of somatic and *de novo* mutations and supporting accurate identification of causative variants. TruSeq Nano DNA kits provide a comprehensive view of the genome, including coding, regulatory, and intronic regions, enabling researchers to access more information from each sequencing run (Figure 6).

Flexible and Inclusive Sample Preparation

The TruSeq family of sample preparation solutions offers several kits for sequencing applications, compatible with a range of research needs and study designs (Table 1). All TruSeq kits support high- and low-throughput studies. The TruSeq Nano DNA kit supports whole-genome sequencing and is ideal for sequencing applications that require sparsely available DNA. These kits provide numerous enhancements to the industry's most widely adopted DNA sample preparation method, empowering all sequencing applications.

Figure 5: Increased Coverage of Challenging Regions



TruSeq Nano DNA libraries demonstrate improved coverage of challenging genomic content. These regions include known human protein coding and non-protein coding exons and genes defined in the RefSeq Genes track in the UCSC Genome Browser.[1] G-Rich regions denote 30 bases with ≥ 80% G. High GC regions are defined as 100 bases with ≥ 75% GC content. Huge GC regions are defined as 100 bases with ≥ 85% GC content. "Difficult" promoters denote the set of 100 promoter regions that are insufficiently covered, which have been empirically defined by the Broad Institute of MIT and Harvard.[2] AT dinucleotides indicate 30 bases of repeated AT dinucleotide.

Efficient Sample Multiplexing

Using a simple procedure, indices are added to sample genomic DNA fragments to provide an innovative solution for sample multiplexing. For the greatest operational efficiency, up to 96 pre-plated, uniquely indexed samples can be pooled and sequenced together in a single flow cell lane on any Illumina sequencing platform. After sequencing, the indices are used to demultiplex the data and accurately assign reads to the proper samples in the pool.

The TruSeq LT kit uses a single index for demultiplexing, while the TruSeq HT kit employs a dual-indexing strategy, using a unique combination of two indices to demultiplex. The LT kit includes up to 24 indices with two sets of 12 each, and the HT kit offers 96 indices.

Streamlined Solution

This inclusive kit contains sample preparation reagents, sample purification beads, and robust TruSeq indices for multiplexing, providing a complete preparation method optimized for the highest performance on all Illumina sequencing platforms. The TruSeq Nano DNA kit leverages the flexibility of two kit options, 24-sample and 96-sample, for scalable experimental design. With a simplified workflow and flexible multiplexing options, the TruSeq Nano DNA protocol offers a streamlined library preparation method that delivers high-quality sequencing data.

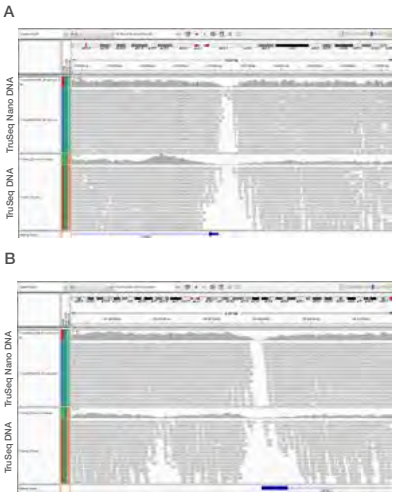
Summary

The TruSeq Nano DNA Sample Preparation Kit optimizes the TruSeq workflow to deliver a low-input sample preparation method for any sequencing application. Low- and high-throughput options and varied insert sizes provide greater flexibility to support a variety of applications and genomic studies. Workflow innovations reduce PCR-induced bias to facilitate detailed and accurate insight into the genome. By leveraging a faster workflow and enhanced data quality, the TruSeq Nano DNA Sample Preparation Kit provides an all-inclusive sample preparation method for genome sequencing applications.

References

1. genome.ucsc.edu
2. www.broadinstitute.org

Figure 6: TruSeq Nano DNA Protocol Reduces Number of Coverage Gaps



Increased coverage of TruSeq Nano DNA libraries results in fewer coverage gaps, demonstrated here in the GC-rich coding regions of the *RNPEPL1* promoter (A) and the *ZBTB34* promoter (B). Sequence information generated by TruSeq Nano DNA prep is shown in the top panels of A and B, while sequence data generated using TruSeq DNA protocol are shown in the lower panels.

Ordering Information

Product	Catalog No.
TruSeq Nano DNA LT Sample Preparation Kit Set A (24 samples)	FC-121-4001
TruSeq Nano DNA LT Sample Preparation Kit Set B (24 samples)	FC-121-4002
TruSeq Nano DNA HT Sample Preparation Kit (96 samples)	FC-121-4003

Illumina • 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. 770-2013-012 Current as of 16 May 2013

Data Sheet: DNA Sequencing

illumina®

Nextera® DNA Sample Preparation Kits

Sequencing's fastest and easiest sample preparation workflow, delivering libraries in 90 minutes.

Highlights

- **Fastest Time to Results**
  Go from DNA to data in less than 8 hours with MiSeq® System
- **Easiest to Use**
  Prepare sequencing-ready samples in 1.5 hours with 15 minutes hands-on time
- **Lowest DNA Input**
  Use just 50 ng DNA per sample, enabling use with samples in limited supply
- **Highest Throughput**
  Index up to 96 samples and use master-mixed reagents to process > 500 samples per week

Figure 1: Nextera DNA Sample Preparation Kit

The Nextera DNA Sample Preparation Kit (96 samples) provides a fast and easy sample preparation workflow, delivering libraries in 90 minutes.

DNA to Data in Record Time

Nextera DNA Sample Preparation Kits provide the fastest and easiest workflow, enabling sequencing-ready libraries to be generated in less than 90 minutes, with less than 15 minutes of hands-on time. DNA is simultaneously fragmented and tagged with sequencing adapters in a single step, using standard lab equipment. Libraries prepared with Nextera kits are compatible with Illumina sequencers (Table 1).

Table 1: Nextera DNA Sample Prep Specifications

Specification	Value
Input DNA	50 ng
Available indexes	Up to 96
Compatible sequencers	HiSeq® NextSeq™, MiSeq, Genome Analyzer IIx, and HiScanSQ Systems
Read lengths supported	Supports all read lengths on any Illumina sequencing system
Typical median insert size	~250 bp
Sample DNA input type	Genomic DNA and PCR amplicons

Breakthrough Chemistry

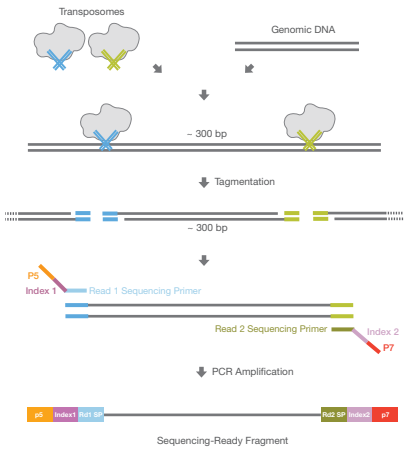
Nextera technology employs a single "tagmentation" reaction to simultaneously fragment and tag DNA with adapters (Figure 2). This process occurs in a single step using master-mixed reagents to provide PCR-ready templates in as little as 15 minutes. Sequencing adapters and indexes are then added to the gDNA fragment by PCR. The optimized Nextera PCR protocol leads to improved performance with GC regions. From start to finish, the complete Nextera sample preparation protocol is over 80% faster than any other method available.

Improved Multiplexing

Nextera DNA Sample Preparation Kits feature an innovative indexing solution for processing and uniquely barcoding up to 96 samples. Multisample studies can be conveniently managed using the Illumina Experiment Manager, a freely available software tool that provides easy reaction setup for plate-based processing.

Following the addition of two indexes to each gDNA fragment, up to 96 uniquely indexed samples can be pooled and sequenced together in a single lane on an Illumina sequencer. After sequencing, the unique combination of the two indexes is used to demultiplex the data and assign reads to the proper sample in the pool. Using this dual barcode approach, Nextera Index Kits only require 20 unique index oligos to process up to 96 samples, providing an easily scalable approach for sample indexing.

Figure 2: Nextera Sample Preparation Biochemistry



Nextera chemistry simultaneously fragments and tags DNA in a single step. A simple PCR amplification then appends sequencing adapters and sample indexes to each fragment.

Accelerated Applications

Nextera DNA Sample Preparation Kits are ideal for experiments where speed and ease are paramount. The low 50 ng DNA input also makes this method amenable to precious samples available in limited quantity. This sample preparation workflow can shorten the overall sequencing workflow time for a wide variety of established applications[1-7] and can be automated for even greater throughput. The combination of the MiSeq System and Nextera DNA Sample Preparation Kits provide rapid DNA to data in as little as 8 hours. These kits enable rapid applications such as small genome and amplicon sequencing, as well as large genome sequencing on any Illumina platform (Table 2).

Summary

The Nextera DNA Sample Preparation Kit provides sequencing's fastest and easiest sample preparation workflow, delivering completed libraries in 90 minutes that are compatible with all Illumina sequencing systems. Nextera enables high-throughput studies with a built-in solution for indexing up to 96 samples with ultra low DNA input. Combined with the MiSeq System, Nextera DNA Sample Preparation Kits enable the fastest DNA to data—all in a single day.

Table 2: Representative Nextera Applications

Examples of Nextera Applications
Large-genome resequencing
Small-genome resequencing
Amplicon resequencing
Clone or plasmid sequencing

References

1. Ramirez MS, Adams MD, Bonomo RA, Centrón D, et al. (2011) Genomic analysis of *Acinetobacter baumannii* A118 by comparison of optical maps: Identification of structures related to its susceptibility phenotype. Antimicrob Agents Chemother, 55(4): 1520–6.
2. Adey A, Morrison HG, Asan, Xun X, Kitzman JO, et al. (2010) Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. Genome Biol 11: R119.
3. Bimber BN, Dudley DM, Lauck M, Becker EA, Chin EN, et al. (2010) Whole-genome characterization of human and simian immunodeficiency virus intrahost diversity by ultradeep pyrosequencing. J Virol 84: 12087–92.
4. Kitzman JO, Mackenzie AP, Adey A, Hiatt JB, Patwardhan RP, et al. (2010) Haplotype-resolved genome sequencing of a Gujarati Indian individual. Nat Biotechnol 29: 59–63.
5. Linnarsson, S. (2010) Recent advances in DNA sequencing methods - General principles of sample preparation. Exp Cell Res 316: 1339–43.
6. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, et al. (2010) Diversity of human copy number variation and multicopy genes. Science 330: 641–646.
7. Voelkerding KV, Dames S, and JD Durtschi (2010) Next generation sequencing for clinical diagnostics-Principles and application to targeted resequencing for hypertrophic cardiomyopathy. J Mol Diagn 12: 539–551.

Ordering Information

Product	Catalog No.
Nextera DNA Sample Preparation Kit (96 samples)	FC-121-1031
Nextera DNA Sample Preparation Kit (24 samples)	FC-121-1030
Nextera Index Kit (96 indexes, 384 samples)	FC-121-1012
Nextera Index Kit (24 indexes, 96 samples)	FC-121-1011
TruSeq Dual Index Sequencing Primer Kit, Single Read (single-use kit)	FC-121-1003
TruSeq Dual Index Sequencing Primer Kit, Paired-End Read (single-use kit)	PE-121-1003

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. 770-2011-021 Current as of 11 March 2014

illumina®

Data Sheet: Sequencing

illumina®

Nextera® XT DNA Sample Preparation Kit

The fastest and easiest sample prep workflow for small genomes, PCR amplicons, and plasmids.

Highlights

- **Rapid Sample Preparation**
  Complete sample prep in as little as 90 minutes with only 15 minutes of hands-on time
- **Fastest Time to Results**
  Go from DNA to data in 8 hours with the MiSeq® System
- **Optimized for Small Genomes, PCR Amplicons, and Plasmids**
  One sample prep kit for many applications
- **Innovative Sample Normalization**
  Eliminates the need for library quantification before sample pooling and sequencing

Introduction

The Nextera XT DNA Sample Preparation Kit enables researchers to prepare sequencing-ready libraries for small genomes (bacteria, archaea, and viruses), PCR amplicons, and plasmids in 90 minutes, with only 15 minutes of hands-on time. The combination of the MiSeq System and Nextera XT DNA Sample Preparation Kits enable you to go from DNA to data in 8 hours (Figure 1). The low amount (1 ng) of input DNA makes this method amenable to precious samples available in limited quantity. Compatible with all Illumina sequencers, Nextera sample preparation can shorten the overall sequencing workflow time for a wide variety of established applications[1-9] and can be automated easily for greater throughput.

Fastest and Easiest Sample Prep Workflow

Using a single "tagmentation" enzymatic reaction, sample DNA is simultaneously fragmented and tagged with adapters. An optimized, limited-cycle PCR protocol amplifies tagged DNA and adds sequencing indexes (Figure 1). From start to finish, the complete Nextera XT protocol is over 80% faster than other available sample preparation methods, and requires the least amount of hands-on time.

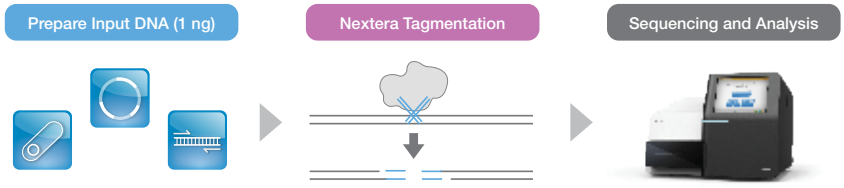
Innovative Sample Normalization

Sample preparation kits for next-generation sequencing result in libraries of varying concentration. To pool samples equally and achieve target cluster densities, time-intensive quantitation methods are often used, followed by dilution and pooling of barcoded samples. The Nextera XT DNA Sample Preparation Kit eliminates the need for library quantification before sample pooling and sequencing by employing a simple bead-based sample normalization step (Figure 2). Prepared libraries are produced at equivalent concentrations enabling pooling by volume—simply pool 5 µl of each library to be sequenced.

Flexible Multiplexing

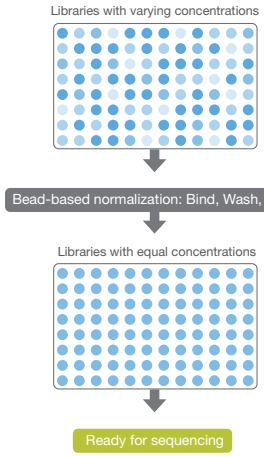
The Nextera XT Sample Preparation Kit features an innovative indexing solution for processing and uniquely barcoding up to 384 samples in a single experiment. Following the addition of two indexes to each DNA fragment, up to 384 uniquely indexed samples can be pooled and sequenced together. After sequencing, the unique combination of the two indexes is used to demultiplex the data and assign reads to the proper sample. Using this dual-barcode approach, Nextera XT Index Kits only require 40 unique index oligos to process up to 384 samples

Figure 1: Nextera XT Sample Preparation Workflow



The combination of Nextera XT and rapid sequencing with the MiSeq System provides a complete DNA to data workflow in only 8 hours.

Figure 2: Innovative Sample Normalization



The Nextera XT Sample Preparation kit eliminates the need for library quantification before sample pooling and sequencing. Libraries of equivalent concentrations are created by employing bead-based sample normalization, as simple as pipetting 5 µl of each library to be sequenced.

for a scalable approach. Multisample studies can be conveniently managed using the Illumina Experiment Manager, a freely available software tool that provides easy reaction setup for plate-based processing.

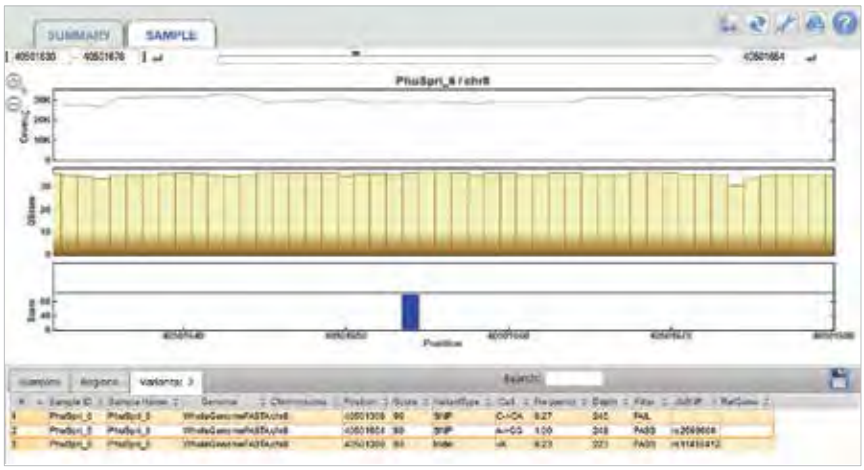
Simple User Interface for Analysis

MiSeq Reporter provides automated on-instrument analysis for various applications including small genome *de novo* or resequencing, PCR amplicon, and plasmid sequencing. Sequencing results and analysis are easy to view and interpret. For example, using the PCR Amplicon workflow in the MiSeq Reporter software, sequence data are automatically categorized into intuitive tabs: Samples, Regions, and Variants (Figure 3). Within each of these tabs, the variant score, quality (Q) score, and sequencing coverage levels can be determined down to single bases, allowing easy analysis of variants of interest.

High Coverage, Accurate Calls

To illustrate the power of amplicon sequencing with Nextera XT and the MiSeq System, nine PCR amplicons of varying sizes were prepared from two different samples of human DNA. Amplicons from each sample were pooled and 1 ng of DNA from each pool was prepared using the Nextera XT kit. Libraries from the two sample pools were combined, sequenced with paired-end 2 × 150 reads on the MiSeq System, and analyzed with MiSeq Reporter using the PCR Amplicon workflow. The approximate mean sequencing coverage values per amplicon and number of variants called (variant score > 99) identified within the amplicons in one of the two samples are shown in Table 1. The output of the MiSeq System supported sequencing of these amplicons to a depth of > 12,000×, enabling

Figure 3: PCR Amplicon Workflow in MiSeq Reporter



MiSeq Reporter provides automated on-instrument analysis for various applications, including PCR amplicon shown here. Samples, regions, and variants are easily accessible, and variant scores, quality (Q) scores, and coverage plots are shown at single nucleotide resolution.

Table 1: Amplicon Coverage and Variants Called

Amplicon Length (bp)	Mean Coverage (thousands of reads)	Variants Called (SNVs/Indels)
953	15.1	4 SNVs
1083	27.4	4 SNVs
1099	22.1	1 SNV
1800	22.4	7 SNVs
1809	17.8	1 SNV
2166	17.6	7 SNVs
3064	12.5	4 SNVs
3064	13.3	1 SNV
3072	14.8 K	1 SNV + 1 indel

confident variant calling. Of the 31 total variants called in this example, 94% are confirmed within the dbSNP database. These results show that coverage is high and even across a range of amplicon sizes, and that variant calls are accurate.

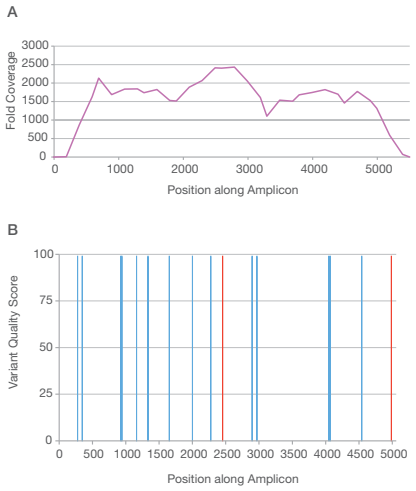
Even Coverage Across Large Amplicons

Large amplicons (> 1 kb) produced by long-range PCR can be easily prepared with the Nextera XT kit and sequenced on any Illumina sequencer. In Figure 4, coverage along amplicon length and position of called variants is shown for a single 5.1 kb amplicon in a highly variable non-coding region of the human genome. The 5.1 kb amplicon was part of a pool of 24 amplicons from human DNA ranging in size from ~300 bp up to 10 kb. Amplicon pools were generated from five different samples, and Nextera XT libraries were made using 1 ng of DNA from each pool. Libraries were combined and single-read sequencing was performed using 1 × 150 bp cycles on MiSeq and analyzed using MiSeq Reporter with the PCR Amplicon workflow.

De Novo Assembly of Small Genomes

To show the utility of Nextera XT for preparing microbial genomes, 1 ng of genomic DNA from *Escherichia coli* reference strain MG1655 was prepared using the Nextera XT kit and sequenced using paired-end 2 × 150 bp reads on the MiSeq System. The data were analyzed using the Assembly workflow on the MiSeq Reporter. Total post-run analysis time for this sample was 28 minutes. Assembly metrics are shown in Table 2. A high-quality assembly was produced, with excellent N50 scores and coverage. This data set is available for analysis in BaseSpace®, the Illumina cloud computing environment[10].

Figure 4: Coverage of Large Amplicons



Panel A: High sequencing coverage (>1,000×) across a 5.1 kb amplicon
Panel B: Within the same amplicon, the position of 16 variants passing filter (14 SNVs in blue + 2 indels in red) is shown, plotted against variant score (a Phred-scaled measure of variant calling accuracy, maximum = 99). Of the 16 variants, 13 are present in dbSNP.

Table 2: *De Novo* Assembly of *E. coli*

Parameter	Value
Percent of genome covered	98%
Number of contigs	314
Maximum contig length	221,108
Base count	4,548,900
N50	111,546
Average coverage per base	184.9

Summary

Nextera XT DNA Sample Preparation Kits are ideal for experiments where speed and ease are of paramount importance. Providing the fastest and easiest sample preparation workflow, the Nextera XT DNA Sample Preparation Kit enables rapid sequencing of small genomes, PCR amplicons, and plasmids. Combined with the MiSeq and NextSeq™ Systems, Nextera XT DNA Sample Preparation Kits enable you to go from DNA to data—all in a single day.

References

1. Loman N, Misra RJ, Dallman TJ, Constantinidou C, Gharbia SE, et al. (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 22 Apr.
2. Gertz J, Varley KE, Davis NS, Baas BJ, Goryshin IY, et al. (2012) Transposase mediated construction of RNA-Seq libraries. Genome Res 22(1): 134–41.
3. Parkinson NJ, Maslau S, Ferneyhough B, Zhang G, Gregory L, et al. (2012) Preparation of high-quality next-generation sequencing libraries from picogram quantities of target DNA. Genome Res 22(1): 125–33.
4. Toprak E, Veres A, Michel J-B, Chait R, Hartl D, et al. (2012) Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat Genet 1(44): 101–106.
5. Raychaudhuri S, Iartchouk O, Chin K, Tan PL, Tai AK, et al. (2011) A rare penetrant mutation in *CFH* confers high risk of age-related macular degeneration. Nat Genet 43(12): 1176–7.
6. Depledge DP, Palser AL, Watson SJ, Lai I Y-C, Gray E, et al. (2011) Specific capture and whole-genome sequencing of viruses from clinical samples. PLoS One 6(11): e27805.
7. Lieberman TD, Michel J-B , Aingaran M, Potter-Bynoe G, Roux D, et al. (2011) Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nat Genet 43(12): 1275–80.
8. Young TS, Walsh CT (2011) Identification of the thiazolyl peptide GE37468 gene cluster from *Streptomyces* ATCC 55365 and heterologous expression in *Streptomyces lividans*. Proc Natl Acad Sci USA 108(32): 13053–8.
9. Adey A, Morrison HG, Asan, Xun X, Kitzman JO, Turner EH, et al. (2010) Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density *in vitro* transposition. Genome Biol 2010;11(12):R119.
10. http://basespace.illumina.com

Nextera XT DNA Sample Prep Kit Specifications

Specification	Value
Sample DNA input type	Genomic DNA, PCR amplicons, plasmids
Input DNA	1 ng
Typical median insert size	< 300 bp
Available indexes	Up to 384
Compatible sequencers	MiSeq, NextSeq, and HiSeq® Systems
Read lengths supported	Supports all read lengths on any Illumina sequencing system

Ordering Information

Product	Catalog No.
Nextera XT DNA Sample Preparation Kit (24 samples)	FC-131-1024
Nextera XT DNA Sample Preparation Kit (96 samples)	FC-131-1096
Nextera XT Index Kit (24 indexes, 96 samples)	FC-131-1001
Nextera XT Index Kit (96 indexes, 384 samples)	FC-131-1002
TruSeq® Dual Index Sequencing Primer Kit, Single Read (single-use kit)*	FC-121-1003
TruSeq Dual Index Sequencing Primer Kit, Paired-End Read (single-use kit)*	PE-121-1003
Nextera XT Index Kit v2, Set A (96 indexes, 384 samples)	FC-131-2001
Nextera XT Index Kit v2, Set B (96 indexes, 384 samples)	FC-131-2002
Nextera XT Index Kit v2, Set C (96 indexes, 384 samples)	FC-131-2003
Nextera XT Index Kit v2, Set D (96 indexes, 384 samples)	FC-131-2004

*Sequencing primer kits are required for all sequencers except the MiSeq System.

Illumina • 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. 770-2012-011 Current as of 11 March 2014

illumina®

Data Sheet: DNA Sequencing

illumina®

Nextera® Mate Pair Sample Preparation Kit

An optimized sample preparation method for long-insert libraries, empowering *de novo* sequencing and structural variant detection.

Highlights

- **Fast and Simple Mate Pair Preparation**
  A simple tagmentation reaction and low DNA input enable library preparation in less than 2 days
- **Dual Protocol Flexibility**
  Gel-free and gel-plus protocols enable a range of applications, including *de novo* assembly and structural variation detection
- **High Data Quality**
  Highly diverse libraries maximize data yield
- **End-to-End Mate Pair Solution**
  Conveniently bundled kit includes reagents and indexes for efficient mate pair preparation

Introduction

Mate pair library preparation generates long-insert paired-end libraries for sequencing. The Nextera Mate Pair Sample Preparation Kit offers two methods, gel-free and gel-plus, to support various applications and input requirements. The robust, low-input, gel-free protocol yields high-diversity libraries that enable deeper sequencing. The size-selection step in the gel-plus protocol generates fragments with a narrow size distribution for structural variation detection. Libraries prepared with the gel-plus protocol also provide sequence information for larger repeat regions, empowering *de novo* genome assembly.

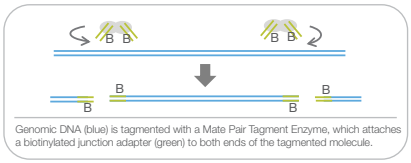
Simplified Mate Pair Workflow

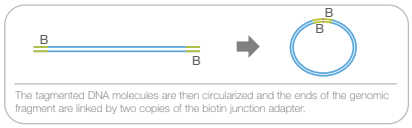
The Nextera Mate Pair protocol provides a simple mate pair workflow for preparing sequencing-ready libraries in less than 2 days (Figure 1). Master-mixed TruSeq® DNA Sample Preparation reagents minimize the number of assay steps, reducing hands-on time to as little as 3 hours. The Nextera "tagmentation" reaction utilizes a specially engineered transposome, the Mate Pair Tagment Enzyme, to simultaneously fragment and tag the DNA sample. This simplified method only biotinylates DNA molecules at the sites of fragmentation, avoiding troublesome internal biotinylation.

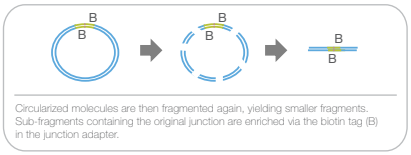
Dual Protocol Flexibility

The flexibility of the Nextera Mate Pair Sample Preparation Kit stems from the availability of two different size-selection options (Table 1). The gel-free protocol, which requires only 1 µg DNA, provides highly diverse mate pair libraries with a broad range of fragment sizes (Figure 2A). This protocol is ideal for routine *de novo* assembly of small bacterial genomes, or for the robust generation of mate pair data for samples with limited DNA. The gel-free protocol offers a faster, simplified option with a lower DNA input requirement to streamline mate pair studies.

Figure 1: Nextera Mate Pair Workflow



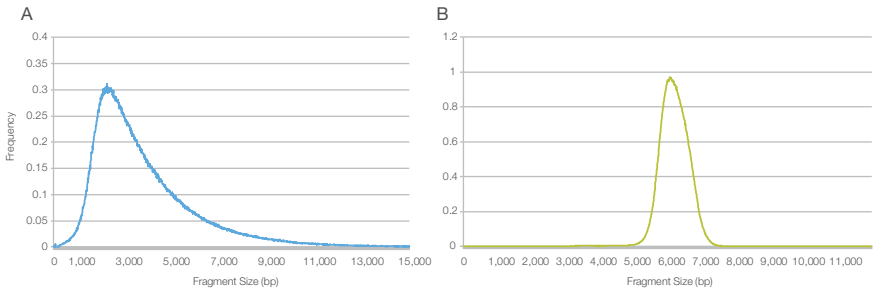




The Nextera Mate Pair Sample Preparation Kit has a simple workflow that enables library preparation in less than 2 days. It supports a range of fragment sizes ~2–12 kb in length, though 2–5 kb fragments are observed at higher frequencies.

The gel-plus protocol, which requires 4 µg DNA and standard agarose gels or Sage Pippin Prep gels[1], offers a more stringent size selection process. The gel-plus protocol produces libraries with narrower size distributions to facilitate structural variation detection (Figure 2B and Figure 3). However, creating gel-plus libraries becomes more difficult as the fragment lengths increase. Greater control over fragment sizes is ideal for more challenging mate pair applications, such as *de novo* assembly of complex genomes and structural variation detection.

Figure 2: Fragment Size Distribution with Dual Protocols



Panel A shows the fragment size distribution of an *E. coli* mate pair library prepared using the Nextera Mate Pair gel-free protocol, resulting in a broad fragment size distribution. Panel B shows the narrow fragment size distribution of an *E. coli* mate pair library generated with the Nextera Mate Pair gel-plus protocol with automated size selection using the Pippin Prep platform.

Highly Diverse Libraries

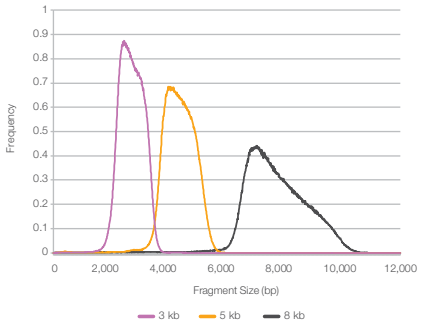
The Nextera tagmentation reaction drives the creation of highly diverse libraries (Table 2) that are compatible with all Illumina sequencing systems. Library diversity is defined as the number of unique fragments in a given library. The Nextera Mate Pair protocol allows for the creation of millions of unique fragments. Such high library diversity generates fewer duplicate reads and yields larger volumes of data.

The Nextera Mate Pair Sample Preparation Kit also provides identifiable junction sequences that mark fragment ends, drastically simplifying data analysis. The presence of searchable junction sequences allows for accurate fragment identification and enables sequencing of longer read lengths, as mate pair junctions can be precisely identified and trimmed accordingly.

Mate Pair Preparation Solution

In addition to Nextera Mate Pair reagents, the comprehensive Nextera kit contains TruSeq DNA sample preparation reagents and indexes. TruSeq on-bead reactions follow the tagmentation and circularization steps (Figure 1), simplifying the purification workflow and reducing sample loss. This integrated solution streamlines the sample preparation workflow, maximizing sequencing efficiency with more samples per lane and enabling rapid multiplexed sequencing of small genomes. The Nextera Mate Pair Sample Preparation Kit is compatible with TruSeq DNA Sample Preparation adapter indexing, supporting 12 indexes per kit for a scalable experimental approach. With all necessary reagents included in one convenient, cost-effective bundle, the Nextera Mate Pair Sample Preparation Kit is an all-in-one solution for fast and simple mate pair library preparation.

Figure 3: Fragment Size Distribution



This figure shows fragment size distributions of three *E. coli* mate pair libraries (3 kb, 5 kb, and 8 kb) created from the same tagmentation reaction. These distributions were generated following the Nextera Mate Pair gel-plus protocol with agarose gel size selection. Though 8 kb fragments are possible with this protocol, 2–5 kb fragments generate libraries with the highest yield and diversity.

Table 1: Nextera Mate Pair Protocols

Protocol	DNA Input	Number of Samples	Size Selections Per Sample	Number of Libraries
Gel-Free	1 µg	48	N/A	48
Gel-Plus with Pippin Prep size selection	4 µg	12	1	12
Gel-Plus with agarose size selection	4 µg	12	Up to 4	Up to 48

Table 2: Nextera Mate Pair Library Diversity*

Preparation	Input DNA	Fragment Size	Diversity†
Nextera Mate Pair Gel-Free	1 µg	~2–12 kb	860 million
Nextera Mate Pair Gel-Plus	4 µg	~2–4 kb	568 million
Nextera Mate Pair Gel-Plus	4 µg	~5–7 kb	396 million
Nextera Mate Pair Gel-Plus	4 µg	~6–10 kb	102 million

* This table demonstrates example diversity values, with diversity reported in number of unique fragments. Actual diversities achieved with this kit may vary and depend on several factors, including DNA input quantity, DNA quality, and precise execution of the protocol.

† Library diversity was calculated from the number of unique read pairs observed in a data set, using a method based on the Lander-Waterman equation[2].

Ordering Information

Product	Catalog No.
Nextera Mate Pair Sample Preparation Kit	FC-132-1001

This kit contains Nextera Mate Pair reagents and TruSeq reagents and indexes.

Summary

With a fast and easy workflow, the Nextera Mate Pair Sample Preparation Kit allows the construction of high-quality sequencing libraries in less than 2 days. The gel-free and gel-plus options provide flexibility for various applications. Transposome-mediated tagmentation, identifiable junction sequences, and indexing capability make the Nextera Mate Pair Sample Preparation Kit a simple and easy solution for mate pair applications.

References

1. www.sagescience.com/products/pippin-prep
2. Lander ES, Waterman MS (1988) Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2: 231–9.

Illumina • 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. 770-2012-052 Current as of 14 March 2014

illumina®

Data Sheet: DNA Sequencing

illumina®

Nextera® Rapid Capture Exomes

A rapid workflow and comprehensive exome content, with unparalleled flexibility.

Highlights

- **Rapid exome preparation and enrichment**
  Prep and enrich 96 exomes in only 1.5 days with less than 5 hours hands-on time
- **Comprehensive exome coverage**
  Two different exome designs are available to access core exonic content or expanded content
- **Kit configurations designed to fit your needs**
  Choose the optimal fit for your system, samples, and study, with more flexible options than ever before
- **Complete support for entire process from sample preparation to sequencing**
  All-in-one kit for prep and enrichment from the world's leading sequencing provider

Overview

Nextera Rapid Capture Exomes are all-in-one kits for sample preparation and exome enrichment that allow researchers to identify coding variants up to 70% faster than other methods. Nextera Rapid Capture Exome delivers 37 Mb of expertly selected exonic content, including challenging regions excluded from other exome designs.

Rapid Exome Prep and Enrichment

Nextera Rapid Capture Exomes provide sample prep and exome enrichment in only 1.5 days. Sequencing with the HiSeq® 2500 or NextSeq™ 500 system enables experiments to go from DNA sample to data in as little as 2.5 days. The speed of Nextera Rapid Capture Exomes enables you to complete projects faster, return results faster, and ultimately publish faster.

Focused Exonic Content

Nextera Rapid Capture Exome has been optimized to provide uniform and specific coverage of 37 Mb of expert-selected exonic content. The probe set was designed to enrich 214,405 exons (Table 1). This focused design, paired with uniform and specific enrichment, enables the most comprehensive exome sequencing available and reliable identification of true, coding variants (Table 2).

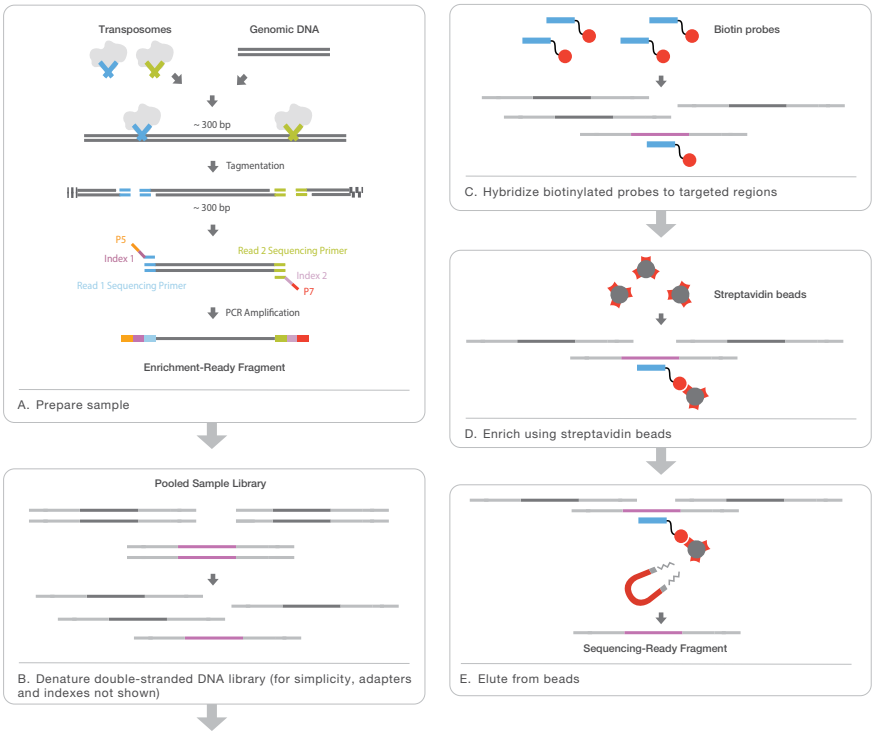
Table 1: Coverage Details

	Nextera Rapid Capture Exome	Nextera Rapid Capture Expanded Exome
Coverage Specifications		
Number of target exons	214,405	201,121
Target content	Coding exons	Exons, UTRs, and miRNA
Percent of Exome Covered (by Database)		
Refseq	98.3%	95.3%
CCDS	98.6%	96.0%
ENSEMBL	97.8%	90.6%
GENCODE v12	98.1%	91.6%

Table 2: Comparison of Rapid Capture Exomes

Specification	Nextera Rapid Capture Exome	Nextera Rapid Capture Expanded Exome
Target size	37 Mb	62 Mb
Genomic DNA input	50 ng	
Hands-on time	5 hours	
Total time	1.5 days	
Batch size	1–96 exomes	

Figure 1: Nextera Rapid Capture Workflow



The Nextera Rapid Capture Exome Kit provides a fast, simple method for isolating the human exome. The streamlined, automation-friendly workflow combines library preparation and exome enrichment steps, and can be completed in 1.5 days with minimum hands-on time.

Greater Coverage with Expanded Exome

Nextera Rapid Capture Expanded Exome features a highly optimized probe set that delivers broad coverage of exons as well as expanded content, such as UTRs and miRNA binding sites. Genome-wide association studies suggest that > 80% of disease-associated variants fall outside coding regions[1]. Analysis of these regions enables researchers to discover variants that affect gene function, at a more affordable price than whole-genome sequencing. The kit includes >340,000 95mer probes, each constructed against the human NCBI37/hg19 reference genome (Table 1). Nextera Rapid Capture Expanded Exome targets a genomic footprint of 62 Mb.

Unmatched Ease

Nextera Rapid Capture Exomes allows researchers to maximize the productivity of their lab personnel and Illumina sequencing technology. The simplicity and speed of the Nextera Rapid Capture assay enables a single technician to prepare and enrich 96 samples in only 1.5 days.

The process starts with rapid Nextera-based sample prep to convert input genomic DNA into adapter-tagged libraries (Figure 1A). This rapid prep requires only 50 ng of input DNA and takes less than 3 hours for a plate of 96 samples. Nextera tagmentation of DNA simultaneously fragments and tags DNA without the need for mechanical shearing.

Table 3: Nextera Rapid Capture Throughput by Illumina Sequencing Systems

Pooling Plexity	Exome Samples per Run				
	MiSeq	NextSeq 500—Mid Output	NextSeq 500—High Output	HiSeq 2500—Rapid Run Mode	HiSeq 2500—High Output
1	Up to 1	–	–	–	–
3	–	Up to 3	–	–	–
6	–	–	Up to 6	Up to 24	Up to 96
9	–	–	Up to 9	Up to 24	Up to 115
12	–	–	Up to 12	Up to 24	Up to 115

Table 3 helps identify which options provide optimal alignment across three vital study design considerations: sequencing instrument, number of exome samples sequenced per run, and the number of exome samples pooled together before enrichment (pooling plexity).

Integrated sample barcodes then allow the pooling of up to 12 samples for a single exome Rapid Capture pull down. Next, libraries are denatured into single-stranded DNA (Figure 1B) and biotin-labeled probes specific to the targeted region are used for the Rapid Capture hybridization (Figure 1C).

The pool is enriched for the desired regions by adding streptavidin beads that bind to the biotinylated probes (Figure 1D). Biotinylated DNA fragments bound to the streptavidin beads are magnetically pulled down from the solution (Figure 1E). The enriched DNA fragments are then eluted from the beads and hybridized for a second Rapid Capture. This entire process is completed in only 1.5 days, enabling a single researcher to efficiently process up to 96 exomes at one time—all without automation.

Summary

Nextera Rapid Capture Exomes provide a fully integrated, rapid solution for exome library prep and enrichment. Available in a wide range of kit configurations (Table 3), as well as two unique designs, Nextera Rapid Capture Exomes provide unparalleled flexibility to optimally align with your specific needs.

References

1. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al. (2009) Finding the missing heritability of complex diseases. Nature 4618: 747–753.

Ordering Information

Kit Description	Catalog No.
Nextera Rapid Capture Exome (8 rxn x 1 plex)	FC-140-1000
Nextera Rapid Capture Exome (8 rxn x 3 plex)	FC-140-1083
Nextera Rapid Capture Exome (8 rxn x 6 plex)	FC-140-1086
Nextera Rapid Capture Exome (8 rxn x 9 plex)	FC-140-1089
Nextera Rapid Capture Exome (2 rxn x 12 plex)	FC-140-1001
Nextera Rapid Capture Exome (4 rxn x 12 plex)	FC-140-1002
Nextera Rapid Capture Exome (8 rxn x 12 plex)	FC-140-1003
Nextera Rapid Capture Expanded Exome (2 rxn x 12 plex)	FC-140-1004
Nextera Rapid Capture Expanded Exome (4 rxn x 12 plex)	FC-140-1005
Nextera Rapid Capture Expanded Exome (8 rxn x 12 plex)	FC-140-1006

Data Sheet: DNA Sequencing

illumina®

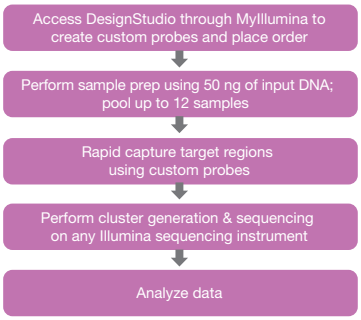
Nextera® Rapid Capture Custom Enrichment

Leverage a superior sample preparation and enrichment workflow for unparalleled access to your regions of interest.

Highlights

- **Integrated sample preparation and enrichment workflow**
  Nextera tagmentation and optimized hybridization reduce workflow duration and generate data faster
- **Target your regions of interest**
  Choose 0.5-15 Mb of custom content, and pool up to 12 samples per enrichment reaction
- **Evolve your design with add-on content**
  Supplement existing panels and keep adding on as your research needs expand

Figure 1: Overview of Nextera Rapid Capture Custom Enrichment



The Nextera Rapid Capture Custom Enrichment Kit is an integral part of a complete and fully supported solution for targeted resequencing.

Introduction

Nextera Rapid Capture Custom Enrichment is an all-in-one assay for sample preparation and custom target enrichment. Nextera tagmentation coupled with optimized target capture ensures the fastest enrichment workflow time for your custom content. The flexible, fully customizable design accommodates up to 15 Mb of custom content so you can focus on the regions of the genome that you care about. The new add-on feature in DesignStudio allows you to iteratively expand your content as new discoveries are made.

Custom Probe Design

The first step in developing any Nextera Rapid Capture Custom Enrichment assay is to design your custom probe set. DesignStudio is a free online user-friendly tool accessed through your MyIllumina account. Designate your regions of interest, refine your custom probe set and place an order for your custom design. DesignStudio uses a complex algorithm to optimize probe set design and alert you to any potential coverage gaps or challenging regions. Desired targets can be added individually or in batches by chromosomal coordinate or gene name.

Unmatched Ease of Workflow

Nextera Rapid Capture Enrichment allows researchers to maximize the productivity of their lab personnel and Illumina sequencing technology. The simplicity and speed of the Nextera Rapid Capture assay enables a single technician to prepare and enrich 12 samples in only 1.5 days.

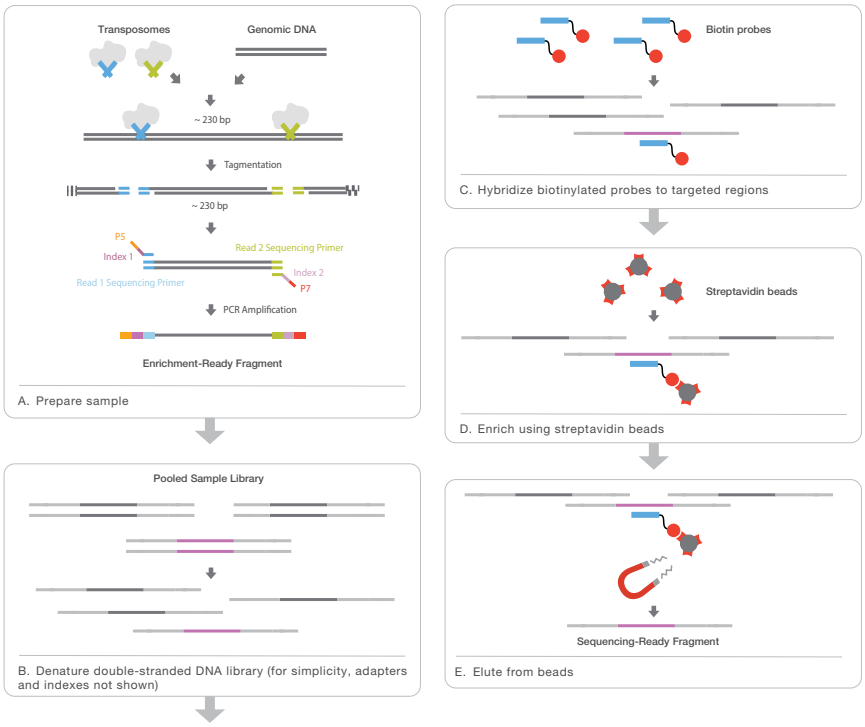
Nextera-based sample preparation generates adapter-tagged libraries from 50 ng input genomic DNA (Figure 2A). Nextera tagmentation of DNA simultaneously fragments and tags DNA without the need for mechanical shearing. Integrated sample barcodes allow the pooling of up to 12 of these adapter ligated sample libraries into a single, hybridization-based, pull down reaction. The pooled libraries are then denatured into single-stranded DNA (Figure 2B) and biotin-labeled probes complementary to the targeted region are used for the Rapid Capture hybridization (Figure 2C). Streptavidin beads are added, which bind to the biotinylated probes that are hybridized to the targeted regions of interest (Figure 2D). Magnetic pull down of the streptavidin beads enriches the targeted regions that are hybridized to biotinylated probes. (Figure 2E). The enriched DNA fragments are then eluted from the beads and a second round of Rapid Capture is completed to increase enrichment specificity. The entire process is completed in only 1.5 days, enabling a single researcher to efficiently process up to 12 samples at one time—all without automation.

Data Analysis

Sequence data generated from custom enrichment samples on HiSeq® and NextSeq™ systems are analyzed using the Enrichment Workflow in the HiSeq Analysis Software (HAS). HAS analysis can be accessed directly via a linux kernel or by using the optional Analysis Visual Controller (AVC) interface[1].

Custom pools sequenced on MiSeq® are analyzed using MiSeq Reporter (MSR). The Enrichment Workflow from both HAS and MSR generates aligned sequence reads in the .bam format using the BWA algorithm and performs indel realignment using the GATK indel realignment tool. Variant calling occurs in the target regions specified in the manifest file. The GATK variant caller generates .vcf

Figure 2: Nextera Rapid Capture Workflow



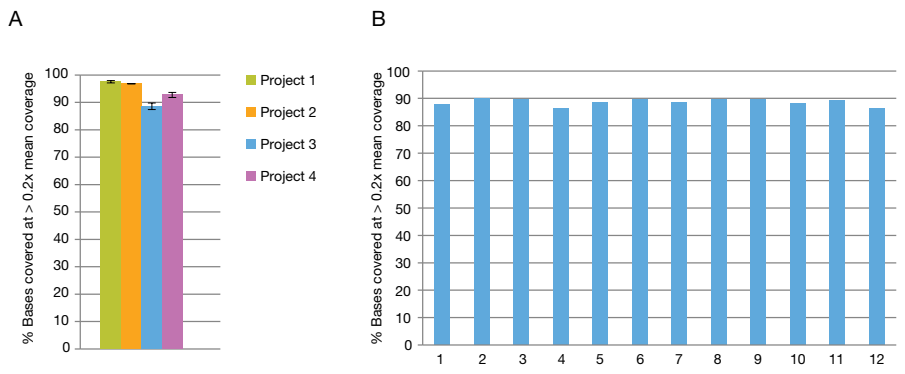
Nextera Rapid Capture Custom Enrichment provides a simple and streamlined in-solution method for isolating and enriching targeted regions of interest. The workflow combines library preparation and exome enrichment steps, and can be completed in 1.5 days with minimum hands-on time.

files that contain genotype, annotation and other information across all sites in the specified target region. Coverage files containing coverage depth in the genome and within gaps is also generated (.CoverageHistogram.txt, .gaps.csv). Additionally, enrichment summary statistics are provided via the.enrichment_summary.csv file or through the CalculateHSMetrics.jar tool within the Picard Suite (.HSmetrics.txt). The enrichment files contain a summary of the on-target and off-target reads/base, average coverage in the target region, % reads that are present at 1×, 10×, 20×, and 50× coverage, read/base enrichment and variant calls information including number of variants (SNP and Indel), Het/Hom and Ts/Tv ratios and the overlap with a standard curated database.

Data Examples

Four different Nextera Rapid Capture Enrichment experiments were performed following the workflow described in Figure 2. Each project included different target regions and coverage depths (Table 1). Representative enrichment and coverage data are shown in Figure 3. In all multiplexed projects, high percent enrichment was achieved, and mean normalized coverage plots show that >85% of bases are covered at 0.2× of the mean coverage. Figure 4 shows that supplementing an existing design (Nextera Rapid Capture Exome) with custom add-on content does not notably decrease coverage uniformity.

Figure 3: High Coverage Uniformity Across Custom 12-plex Pools



Nextera Rapid Capture Custom Enrichment provides uniform target enrichment across different custom probe sets and individual samples within a 12-plex pool. A. Coverage uniformity is shown as % of targeted bases that are represented by >0.2× mean coverage. Mean coverage for these custom probe sets can be found in Table 1. Error bars show SD of uniformity across the 12 pooled samples for each project. B. Coverage uniformity for each of 12 pooled samples within Project 3 is shown. Mean coverage for this run was 300×, and % of targeted bases that were covered at > 60× are shown.

Table 1: Sequencing Details for Example Projects

Project	Content	Mean Coverage	% On Target Bases*
1**	0.5 Mb	1500×	88.6
2**	0.5 Mb	146×	79.5
3†	3.5 Mb	300×	80.1
4**	7 Mb	152×	72.5

*Calculated using Picard Hybrid Selection tool with 250 bp padding[2]

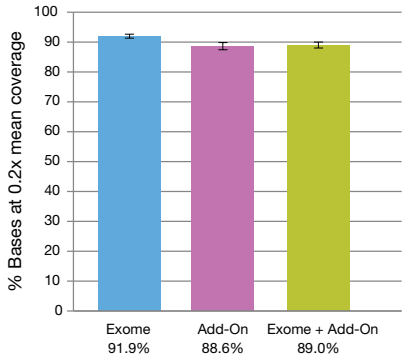
**Sequenced on HiSeq

†Sequenced on MiSeq

Summary

Nextera Rapid Capture Custom Enrichment leverages a superior integrated sample prep and enrichment workflow to provide unparalleled access to your genomic regions of interest. Not only will you be able to perform targeted sequencing using only 50 ng of input DNA, you'll do so faster and more efficiently than ever before. Take advantage of robust add-on functionality to refine your content over time, or add regions of unique interest to established panels such as Nextera Rapid Capture Exome or other TruSight™ content sets.

Figure 4: Add-On Content Retains High Coverage



High coverage uniformity is maintained when 3.5 Mb of add-on content is added to the Nextera Rapid Capture Exome. All samples were run as 12-plex pools.

Nextera Rapid Capture Custom Enrichment Details

Enrichment Efficiency*	>70%
Coverage Uniformity (0.2x mean)	>85%
Content Range	0.5–15 Mb
Samples in Pre-Enrichment Pooling	Up to 12
Sample Input	50 ng
Library Insert Size	230

*Target values will vary due to custom designs.

Ordering Information

Product	Catalog No.
Nextera Rapid Capture Custom (48 samples) Compatible with designs of 3,000-10,000 custom enrichment probes	FC-140-1007
Nextera Rapid Capture Custom (96 samples) Compatible with designs of 3,000-10,000 custom enrichment probes	FC-140-1008
Nextera Rapid Capture Custom (288 samples) Compatible with designs of 3,000-67,000 custom enrichment probes	FC-140-1009

Learn More

To learn more about complete solutions for targeted resequencing, visit www.illumina.com/applications/sequencing/targeted_resequencing.ilmn.

References

1. http://support.illumina.com/sequencing/sequencing_software/analysis_visual_controller_avc.ilmn
2. http://picard.sourceforge.net

Illumina • 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. 770-2013-025 Current as of 21 December 2013

illumina®

epicentre®
an Illumina® company

Methyl-Seq

EpiGnome™ Methyl-Seq Kit

Unlock limited samples (50-100ng DNA input) to discover methylation patterns of all CpG, CHH & CHG regions.

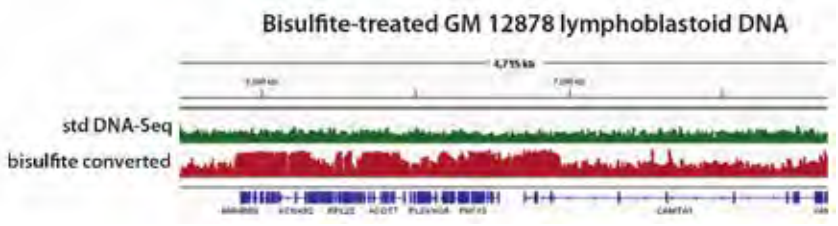
- Unlock small samples (50-100ng DNA input)
- Pre-library bisulfite conversion
- Comprehensive, whole genome results
- 5 hour method
- Informatics app note demystifies analysis
- ***Capture full sample diversity***

Sequence the entire sample–no loss of information!

The process of bisulfite treatment denatures genomic DNA into single stranded DNA. EpiGnome converts single stranded DNA into an Illumina® sequencing library. All ssDNA fragments are captured into an Illumina sequencing library during the EpiGnome procedure, therefore eliminating sample loss associated with other methods.

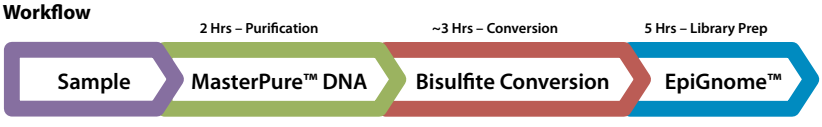
Calico cats are domestic cats with a spotted or parti-colored coat that is predominantly white, with patches of two other colors. Calico cats are almost always female because the X chromosome determines the coat color. During embryonic development, one X chromosome is hypermethylated and inactivated. The remaining X chromosome determines coat color.

Figure 1. EpiGnome is sensitive to CpG methylation patterns.



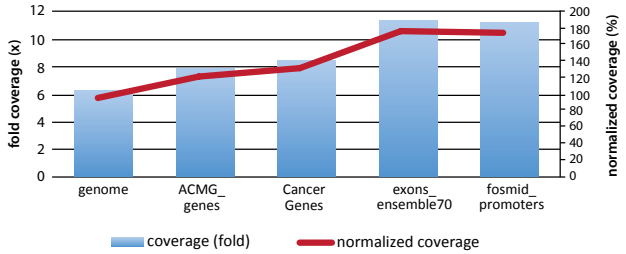
CpG methylation patterns across region of chr1 show variable CpG methylation (red) from 50 ng input of GM12878 lymphoblastoid gDNA treated with bisulfite. Comparison to coverage patterns from non-bisulfite treated (green) gDNA shows the methylated regions of chromosome 1.

Workflow



www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 009a

Figure 2. Deep coverage of genes of interest.



ACMG_genes: designated as medically relevant.

Cancer genes: protein coding genes known to be involved in cancer.

fosmid_promoters: of high interest and difficult to sequence.

EpiGnome WGBS method yields high coverage of genes of interest for Cancer genes and those that have been defined as medically relevant by the American College of Medical Genetics.

Deep coverage of critical genomic regions

Depth of coverage is enhanced in genomic areas with biological utility (Figure 2). EpiGnome captures full sample diversity of critical areas including:

- Coding region start and end for exons from the canonical transcript of protein coding genes for genes known to be involved in cancer, taken from SOMA and CRUK panels as well as literature derived Cancer genes.
- Genes defined by the American College of Medical Genetics as being medically relevant (ACMG_genes)
- Exonic coding regions from Ensemble 70 (exons_ ensemble70)
- List of 100 promoters defined by the Broad Institute as being of high interest and difficult to sequence (fosmid_ promoters)

Coverage was obtained from 125.4 million reads in a single lane of a HiSeq. Increasing throughput of the HiSeq Systems enables complete methylation information to be captured from a growing number of samples.

Cat. #	Quantity
EpiGnome™ Methyl-Seq Kit	
EGMK81312	12 reactions
EGMK91324	24 reactions
EGMK91396	96 reactions
EpiGnome™ Index PCR Primers	
EGIDX81312	12 indexes, 10 reactions each
FailSafe™ PCR Enzyme Mix	
FSE51100	100 units

The FailSafe™ PCR Enzyme Mix is required for EpiGnome Methyl-Seq Kit.

Success begins with purification

MasterPure™ DNA Purification Kit

Purification is an important step to prepare your sample. MasterPure safely removes unwanted material to give you pure DNA.

MasterPure offers unique benefits:

- Very high yields
- Recover <90% of theoretical yield
- Safe and nontoxic
- Available for all sample sizes

Cat. #	Quantity
MasterPure™ Complete DNA and RNA Purification Kit	
MC85200	200 Purifications
MC89010	10 Purifications

Data Sheet: Sequencing

illumina®

TruSeq® ChIP Sample 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.

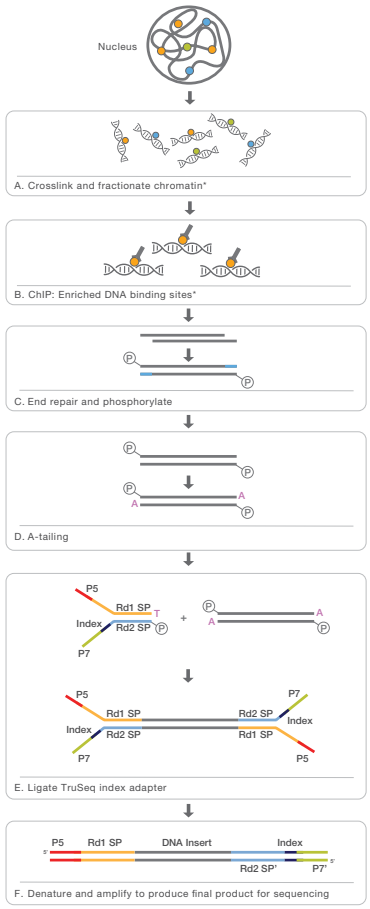
As the output of NGS systems has increased, ChIP-Seq researchers increasingly require a combination of highly multiplexed sequencing and simple, streamlined workflows. TruSeq ChIP Sample Preparation Kits meet those demands, offering a simple, cost-effective solution for obtaining visibility into the 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 sample 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 Sample 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).

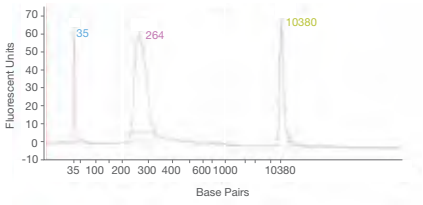
*Steps A and B are performed prior to the TruSeq ChIP Sample Prep workflow.

Figure 1: ChIP-Seq Workflow



The simple, streamlined TruSeq ChIP Sample 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 Sample 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 Sample 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 Sample Preparation Kits Compared to ENCODE Reference Peak Data

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

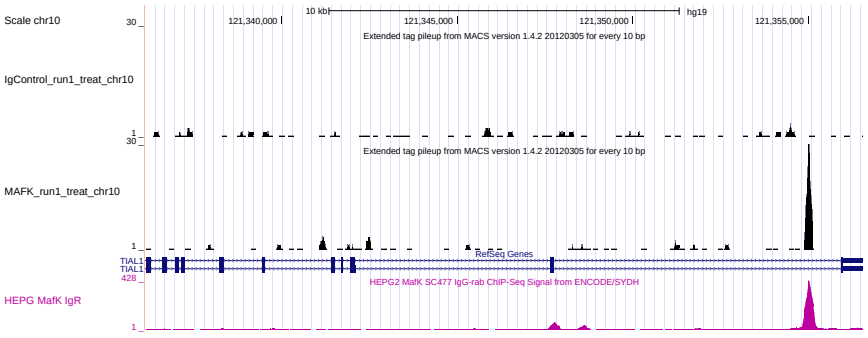
TruSeq Data Quality

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 Sample 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 Sample 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 Sample 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

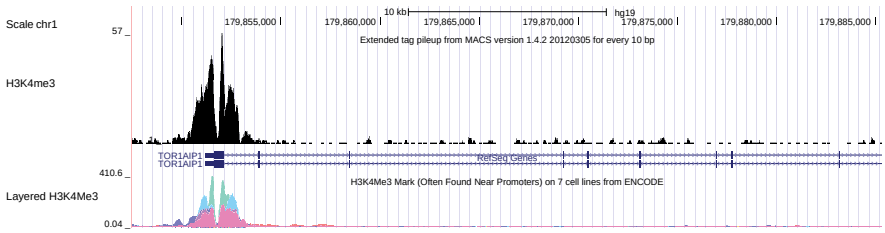
Using the TruSeq ChIP Sample Preparation Kit, a library was generated for transcription factor MafK using 5 ng of input DNA (Figure 2) derived from a ChIP performed in HELA cells. Sequencing data were generated using a single MiSeq run. Quality-filtered, BAM output files were then entered into the MACS 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 Sample 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 Sample 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 Sample 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 Sample Preparation Kit, Set A (12 indexes, 48 samples)	IP-202-1012
TruSeq ChIP Sample 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 22 August 2012

DNA-Sequencing	
Description	**Catalog Number**
MasterPure™ Complete DNA and RNA Purification Kit	MC85200
MasterPure™ DNA Purification Kit	MCD85201
TruSeq DNA PCR-Free LT Sample Preparation Kit - Set A	FC-121-3001
TruSeq DNA PCR-Free LT Sample Preparation Kit - Set B	FC-121-3002
TruSeq DNA PCR-Free HT Sample Preparation Kit	FC-121-3003
TruSeq Nano DNA LT Sample Preparation Kit - Set A	FC-121-4001
TruSeq Nano DNA LT Sample Preparation Kit - Set B	FC-121-4002
TruSeq Nano DNA HT Sample Preparation Kit	FC-121-4003
Nextera Rapid Capture Exome (8 rxn x 1 Plex)	FC-140-1000
Nextera Rapid Capture Exome (8 rxn x 3 Plex)	FC-140-1083
Nextera Rapid Capture Exome (8 rxn x 6 Plex)	FC-140-1086
Nextera Rapid Capture Exome (8 rxn x 9 Plex)	FC-140-1089
Nextera Rapid Capture Exome (2 rxn x 12 Plex)	FC-140-1001
Nextera Rapid Capture Exome (4 rxn x 12 Plex)	FC-140-1002
Nextera Rapid Capture Exome (8 rxn x 12 Plex)	FC-140-1003
Nextera Rapid Capture Expanded Exome (2 rxn x 12 Plex)	FC-140-1004
Nextera Rapid Capture Expanded Exome (4 rxn x 12 Plex)	FC-140-1005
Nextera Rapid Capture Expanded Exome (8 rxn x 12 Plex)	FC-140-1006
EpiGnome™ Methyl-Seq Kit	EGMK81312

ChIP	
Description	**Catalog Number**
TruSeq ChIP Sample Preparation Kit - Set A	IP-202-1012
TruSeq ChIP Sample Preparation Kit - Set B	IP-202-1024

Methylation Arrays	
Description	**Catalog Number**
HumanMethylation450 DNA Analysis BeadChip Kit (24 samples)	WG-314-1003
HumanMethylation450 DNA Analysis BeadChip Kit (48 samples)	WG-314-1001
HumanMethylation450 DNA Analysis BeadChip Kit (96 samples)	WG-314-1002

RNA-sequencing	
Description	**Catalog Number**
MasterPure™ Complete DNA and RNA Purification Kit	MC85200
TotalScript™ RNA-Seq Kit	TSRNA 12924
ScriptSeq™ Complete Gold Kit (Blood)	BGGB1306
ScriptSeq™ Complete Gold Kit (Blood) - Low Input	SCL24GBL
Ribo-Zero Magnetic Gold Kit (Yeast)	MRZY1324
ScriptSeq™ Complete Gold Kit (Yeast)	BGY1324
ScriptSeq™ Complete Gold Kit (Yeast) - Low Input	SCGL6Y
ARTseq™ Ribosome Profiling Kit - Mammalian	RPHMR12126
ARTseq™ Ribosome Profiling Kit - Yeast	RPYSC12116
Any species	
TruSeq® Stranded mRNA LT Set A	RS-122-2101
TruSeq® Stranded mRNA LT - Set B	RS-122-2102
TruSeq® Stranded mRNA HT	RS-122-2103
TruSeq™ RNA Sample Prep Kit v2 -Set A (48rxn)	RS-122-2001
TruSeq™ RNA Sample Prep Kit v2 -Set B (48rxn)	RS-122-2002
Human/Mouse/Rat	
TruSeq® Strnd Total RNA LT(w/Ribo-Zero™ Human/Mouse/Rat)Set A	RS-122-2201
TruSeq® Strnd Total RNA LT(w/Ribo-Zero™ Human/Mouse/Rat)Set B	RS-122-2202
TruSeq® StrndTotal RNA HT (w/ Ribo-Zero™ Human/Mouse/Rat)	RS-122-2203
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Gold) Set A	RS-122-2301
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Gold) Set B	RS-122-2302
TruSeq® Stranded Total RNA HT (w/ Ribo-Zero™ Gold)	RS-122-2303
Human/Mouse/Rat (Blood-derived)	
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Globin) Set A	RS-122-2501
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Globin) Set B	RS-122-2502
TruSeq® Stranded Total RNA HT (w/ Ribo-Zero™ Globin)	RS-122-2503
Plant	
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Plant) Set A	RS-122-2401
TruSeq® Stranded Total RNA LT (w/ Ribo-Zero™ Plant) Set B	RS-122-2402
TruSeq® Stranded Total RNA HT (w/ Ribo-Zero™ Plant)	RS-122-2403

Small RNA-sequencing	
Description	**Catalog Number**
TruSeq® Small RNA Sample Prep Kit -Set A	RS-200-0012
TruSeq® Small RNA Sample Prep Kit -Set B	RS-200-0024
TruSeq® Small RNA Sample Prep Kit -Set C	RS-200-0036
TruSeq® Small RNA Sample Prep Kit -Set D	RS-200-0048

Targeted RNA-Sequencing	
Description	**Catalog Number**
TruSeq Targeted RNA Expression Custom Components	
TruSeq Targeted RNA Custom Kit (48 Samples)	RT-101-1001
TruSeq Targeted RNA Custom Kit (96 Samples)	RT-102-1001
TruSeq Targeted RNA supplemental content (48 Samples)	RT-801-1001
TruSeq Targeted RNA supplemental content (96 Samples)	RT-802-1001
TruSeq Targeted RNA Index Kit	RT-401-1001
TruSeq Targeted RNA Expression Fixed Panels	
TruSeq Targeted RNA Apoptosis Panel Kit (48 Samples)	RT-201-1010
TruSeq Targeted RNA Apoptosis Panel Kit (96 Samples)	RT-202-1010
TruSeq Targeted RNA Cardiotoxicity Panel Kit (48 Samples)	RT-201-1009
TruSeq Targeted RNA Cardiotoxicity Panel Kit (96 Samples)	RT-202-1009
TruSeq Targeted RNA Cell Cycle Panel Kit (48 Samples)	RT-201-1003
TruSeq Targeted RNA Cell Cycle Panel Kit (96 Samples)	RT-202-1003
TruSeq Targeted RNA Cytochrome p450 Panel Kit (48 Samples)	RT-201-1006
TruSeq Targeted RNA Cytochrome p450 Panel Kit (96 Samples)	RT-202-1006
TruSeq Targeted RNA HedgeHog Panel Kit (48 Samples)	RT-201-1002
TruSeq Targeted RNA HedgeHog Panel Kit (96 Samples)	RT-202-1002
TruSeq Targeted RNA Neurodegeneration Panel Kit (48 Samples)	RT-201-1001
TruSeq Targeted RNA Neurodegeneration Panel Kit (96 Samples)	RT-202-1001
TruSeq Targeted RNA NFkB Panel Kit (48 Samples)	RT-201-1008
TruSeq Targeted RNA NFkB Panel Kit (96 Samples)	RT-202-1008
TruSeq Targeted RNA Stem Cell Panel Kit (48 Samples)	RT-201-1005
TruSeq Targeted RNA Stem Cell Panel Kit (96 Samples)	RT-202-1005
TruSeq Targeted RNA TP53 Pathway Panel Kit (48 Samples)	RT-201-1007
TruSeq Targeted RNA TP53 Pathway Panel Kit (96 Samples)	RT-202-1007
TruSeq Targeted RNA Wnt Pathway Panel Kit (48 Samples)	RT-201-1004
TruSeq Targeted RNA Wnt Pathway Panel Kit (96 Samples)	RT-202-1004

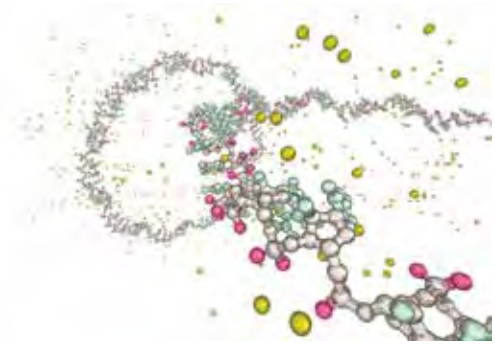
epicentre®
an Illumina® company

RNA-Seq without rRNA Depletion

TotalScript™ RNA-Seq Kits

- Powered by Nextera™
- 1-5 ng Input RNA
- Designed for precious samples
- Rapid method with 5 hr workflow

- rRNA removal not required, begin with total RNA
- Directional libraries
- 12 Indexes available

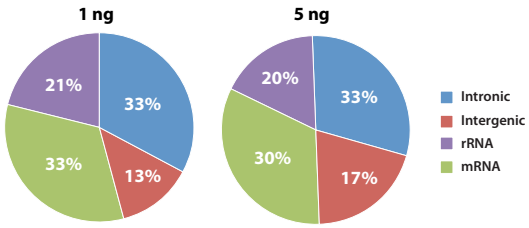


Powered by Nextera! TotalScript RNA-Seq Kit is designed for RNA-Seq of precious samples, and only 1-5 ng of intact total RNA is needed for each sample. No need for poly(A) enrichment or rRNA removal. Sequencing data is similar to data from libraries using much more RNA.

New kinds of samples can now be sequenced, including:

- Cancer samples
- Stem cells
- Other low input samples

Figure 1. Consistent results from TotalScript.



Retain more sample

Prevent transcript loss

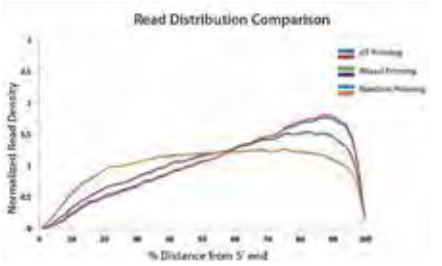
TotalScript produces consistent results from small amounts of sample (Fig. 1). 1 ng or 5 ng of total RNA was prepared with TotalScript. Results show similiar amounts of coding and non-coding coverage between samples. The sample was Universal Human Reference RNA (UHR) total RNA.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 007

Figure 2. You choose the coverage profile.

Method	Total RNA Input (ng)	% rRNA	Coverage
Random Priming	1-5	<40%	Even
Mixed Priming	1-5	<25%	Slight 3′ Bias
dT Priming	1-5	<5%	3′ Bias



You choose the desired rRNA content and transcript coverage with TotalScript™

Three options are included in every TotalScript kit (Fig. 2). All options produce directional libraries from very small amounts of total RNA.

1. Random Hexamer Primer option produces even transcript coverage with <40% of reads mapping to rRNA.
2. Mixed Primer option produces good transcript coverage with <25% rRNA mapped reads.
3. Oligo(dT) Primer option produces <5% rRNA reads with transcript coverage strongest at the 3′ end.

Different sources of RNA may produce different levels of rRNA contamination.

TotalScript RNA-Seq libraries shown were made from 5 ng of total UHR RNA using the Optimized Buffer included with TotalScript (Fig 2).

Cat. #	Quantity
TotalScript™ RNA-Seq Kit	
TSRNA12924	24 Reactions
TSRNA1296	12 Reactions
TotalScript™ Index Kit	
TSIDX12910	11 indexes

Success begins with purification

MasterPure™ RNA Purification Kit

Purification is an important step to prepare your sample. MasterPure safely removes unwanted material to give you pure, intact total RNA.

MasterPure offers unique benefits:

- Keep RNA intact (does not degrade RNA)
- Retain RNA diversity (including small RNA)
- Maximize genes discovered
- Available for all sample sizes

Cat. #	Quantity
MasterPure™ RNA Purification Kit (for isolating RNA only)	
MCR85102	100 Purifications

Data Sheet: Illumina® Sequencing

illumina®

TruSeq™ RNA and DNA Sample Preparation Kits v2

Master-mixed reagents, optimized adapter design, and a flexible workflow provide a simple, cost-effective method for preparing RNA and DNA samples for scalable next-generation sequencing.

Highlights

- **Simple Workflow for RNA and DNA:** Master-mixed reagents and minimal hands-on steps.
- **Scalable and Cost-Effective Solution:** Optimized formulations and plate-based processing enables large-scale studies at a lower cost.
- **Enhanced Multiplex Performance:** Twenty-four adaptor-embedded indexes enable high-throughput processing and greater application flexibility.
- **High-Throughput Gene Expression Studies:** Gel-free, automation-friendly RNA sample preparation for rapid expression profiling.

Introduction

Illumina next-generation sequencing (NGS) technologies continue to evolve, offering increasingly higher output in less time. Keeping pace with these developments requires improvements in sample preparation. To maximize the benefits of NGS and enable delivery of the highest data accuracy, Illumina offers the TruSeq RNA and DNA Sample Preparation Kits (Figure 1).

The TruSeq RNA and DNA Sample Preparation Kits provide a simple, cost-effective solution for generating libraries from total RNA or genomic DNA that are compatible with Illumina's unparalleled sequencing output. Master-mixed reagents eliminate the majority of pipetting steps and reduce the amount of clean-up, as compared to previous methods, minimizing hands-on time. New automation-friendly workflow formats enable parallel processing of up to 96 samples. This results in economical, high-throughput RNA or DNA sequencing studies achieved with the easiest-to-use sample preparation workflow offered by any NGS platform.

Simple and Cost-Effective Solution

Whether processing samples for RNA-Seq, genomic sequencing, or exome enrichment, the TruSeq kits provide significantly improved library preparation over previously used methods. New protocols reduce the number of purification, sample transfer, and pipetting steps. The new universal, methylated adaptor design incorporates an index sequence at the initial ligation step for improved workflow efficiency and more robust multiplex sequencing. For maximum flexibility, the same TruSeq kit can be used to prepare samples for single-read, paired-end, and multiplexed sequencing on all Illumina sequencing instruments.

TruSeq DNA and RNA Sample Prep kits include gel-free protocols that eliminate the time-intensive gel purification step found in other methods, making the process more consistent and fully automatable.

The gel-free protocol for TruSeq DNA sample preparation is available for target enrichment using the TruSeq Exome Enrichment or TruSeq Custom Enrichment kits.

TruSeq sample preparation makes RNA sequencing for high-throughput experiments more affordable, enabling gene expression profiling studies to be performed with NGS at a lower cost than arrays. It also provides a cost-effective DNA sequencing solution for large-scale whole-genome resequencing, targeted resequencing, *de novo* sequencing, metagenomics, and methlyation studies.

Enhanced Multiplex Performance

TruSeq kits take advantage of improved multiplexing capabilities to increase throughput and consistency, without compromising results. Both the RNA and DNA preparation kits include adapters containing unique index sequences that are ligated to sample fragments at the beginning of the library construction process. This allows the samples to be pooled and then individually identified during downstream analysis. The result is a more efficient, streamlined workflow that leads directly into a superior multiplexing solution. There are no additional PCR steps required for index incorporation, enabling a robust, easy-to-follow procedure. With 24 unique indexes available, up to 384 samples can be processed in parallel on a single HiSeq 2000 run.

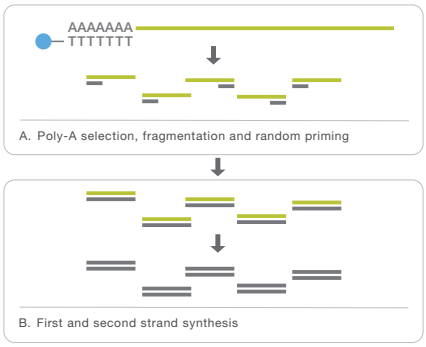
TruSeq RNA Sample Preparation

With TruSeq reagents, researchers can quickly and easily prepare samples for next-generation sequencing (Figure 2). Improvements in the RNA to cDNA conversion steps have significantly enhanced the overall workflow and performance of the assay (Figure 3).

Figure 1: TruSeq Sample Preparation Kits

TruSeq Sample Preparation Kits are available for both genomic DNA and RNA samples.

Figure 2: Optimized TruSeq RNA Sample Preparation



Starting with total RNA, mRNA is polyA-selected and fragmented. It then undergoes first- and second-strand synthesis to produce products ready for library construction (Figure 4).

Starting with total RNA, the messenger RNA is first purified using polyA selection (Figure 2A), then chemically fragmented and converted into single-stranded cDNA using random hexamer priming. Next, the second strand is generated to create double-stranded cDNA (Figure 2B) that is ready for the TruSeq library construction workflow (Figure 4).

Efficiencies gained in the polyA selection process, including reduced sample transfers, removal of precipitation steps, and combining of elution and fragmentation into a single step, enable parallel processing of up to 48 samples in approximately one hour. This represents a 75% reduction in hands-on time for this portion of library construction. Improving performance, the optimized random hexamer priming strategy provides the most even coverage across transcripts, while allowing user-defined adjustments for longer or shorter insert lengths.

Eliminating all column purification and gel selection steps from the workflow removes the most time-intensive portions, while improving the assay robustness. It also allows for decreased input levels of RNA—as low as 100 ng— and maintains single copy per gene sensitivity.

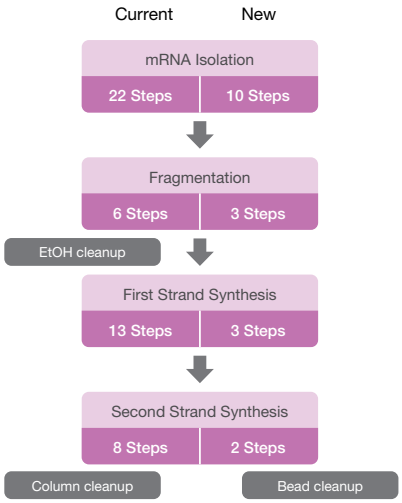
TruSeq DNA Sample Preparation

The TruSeq DNA Sample Preparation Kits are used to prepare DNA libraries with insert sizes from 300–500 bp for single, paired-end, and multiplexed sequencing. The protocol supports shearing by either sonication or nebulization with a low input requirement of 1 ug of DNA.

Sequence-Ready Libraries

Library construction begins with either double-stranded cDNA synthesized from RNA or fragmented gDNA (Figure 4A). Blunt-end DNA fragments are generated using a combination of fill-in reactions and exonuclease activity (Figure 4B). An 'A'- base is then added to the blunt ends of each strand, preparing them for ligation to the sequencing adapters (Figures 4C). Each adapter contains a 'T'-base overhang on 3'-end providing a complementary overhang for ligating the adapter

Figure 3: TruSeq RNA Sample Preparation Reagents Provide Significant Savings in Time and Effort



	Current Methods	TruSeq Methods	Savings
No. of Steps	49	18	31
Time (hours)	16	12	25%

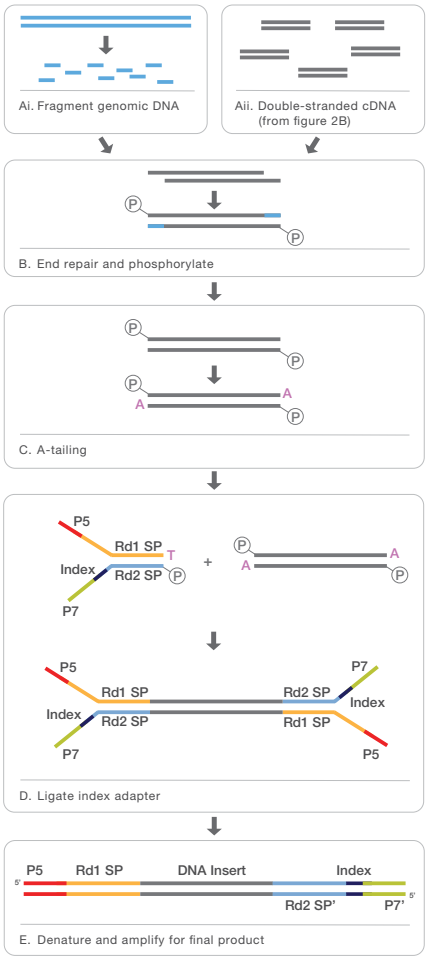
Compared to current methods for preparing mRNA samples for sequencing, use of the TruSeq reagents significantly reduces the number of steps and hands-on time.

Table 1: Savings When Processing 96 Samples

> 50% of pipetting steps eliminated
> 50% of reagent tubes eliminated
> 75% of clean-up steps eliminated
> 50% of sample transfer steps eliminated

Compared to previous kits, processing multiple samples with the new TruSeq Sample Preparation Kits provides significant reductions in library construction costs, the number of steps, hands-on time, and PCR dependency.

Figure 4: Adapter Ligation Results in Sequence-Ready Constructs without PCR



Library construction begins with either fragmented genomic DNA or double-stranded cDNA produced from total RNA (Figure 4A). Blunt-end fragments are created (Figure 4B) and an A-base is then added (Figure 4C) to prepare for indexed adapter ligation (Figure 4D). Final product is created (Figure 4E), which is ready for amplification on either the cBot or the Cluster Station.

to the A-tailed fragmented DNA. These newly redesigned adapters contain the full complement of sequencing primer hybridization sites for single, paired-end, and multiplexed reads. This eliminates the need for additional PCR steps to add the index tag and multiplex primer sites (Figure 4D). Following the denaturation and amplification steps (Figure 4E), libraries can be pooled with up to 12 samples per lane (96 sample per flow cell) for cluster generation on either cBot or the Cluster Station.

Master-mixed reagents and an optimized protocol improve the library construction workflow, significantly decreasing hands-on time and reducing the number of clean-up steps when processing samples for large-scale studies (Table 1). The simple and scalable workflow allows for high-throughput and automation-friendly solutions, as well as simultaneous manual processing for up to 96 samples. In addition, enhanced troubleshooting features are incorporated into each step of the workflow, with quality control sequences supported by Illumina RTA software.

Enhanced Quality Controls

Specific Quality Control (QC) sequences, consisting of double-stranded DNA fragments, are present in each enzymatic reaction of the TruSeq sample preparation protocol: end repair, A-tailing, and ligation. During analysis, the QC sequences are recognized by the RTA software (versions 1.8 and later) and isolated from the sample data. The presence of these controls indicates that its corresponding step was successful. If a step was unsuccessful, the control sequences will be substantially reduced. QC controls assist in comparison between experiments and greatly facilitate troubleshooting.

Designed For Automation

The TruSeq Sample Preparation Kits are compatible with high-throughput, automated processing workflows. Sample preparation can be performed in standard 96-well microplates with master-mixed reagent pipetting volumes optimized for liquid-handling robots. Barcodes on reagents and plates allow end-to-end sample tracking and ensure that the correct reagents are used for the correct protocol, mitigating potential tracking errors.

Part of an Integrated Sequencing Solution

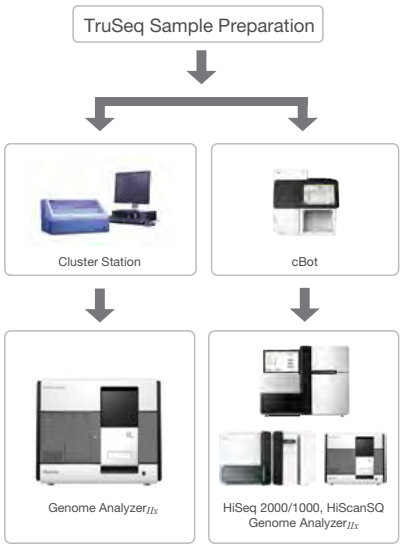
Samples processed with the TruSeq Sample Preparation Kits can be amplified on either the cBot Automated Cluster Generation System or the Cluster Station and used with any of Illumina's next-generation sequencing instruments, including HiSeq™ 2000, HiSeq 1000, HiScan™SQ, Genome Analyzer$_{IIx}$ (Figure 5).

Summary

Illumina's new TruSeq Sample Preparation Kits enable simplicity, convenience, and affordability for library preparation. Enhanced multiplexing with 24 unique indexes allows efficient high-throughput processing. The pre-configured reagents, streamlined workflow, and automation-friendly protocol save researchers time and effort in their next-generation sequencing pursuits, ultimately leading to faster discovery and publication.

Learn more about Illumina's next-generation sequencing solutions at www.illumina.com/sequencing.

Figure 5. Illumina's Complete Sequencing Solution



The TruSeq Sample Preparation Kits readily fit in with Illumina's advanced next-generation sequencing solutions.

Ordering Information

Product	Catalog No.
For RNA Preparation	
TruSeq RNA Sample Preparation Kit v2, Set A (12 indexes, 48 samples)	RS-122-2001
TruSeq RNA Sample Preparation Kit v2, Set B (12 indexes, 48 samples)	RS-122-2002
For DNA Preparation	
TruSeq DNA Sample Preparation Kit v2, Set A (12 indexes, 48 samples)	FC-121-2001
TruSeq DNA Sample Preparation Kit v2, Set B (12 indexes, 48 samples)	FC-121-2002
For Cluster Generation on cBot and Sequencing on the HiSeq 2000/1000 and HiScanSQ	
TruSeq Paired-End Cluster Kit v3—cBot—HS (1 flow cell)	PE-401-3001
TruSeq Single-Read Cluster Kit v3—cBot—HS (1 flow cell)	GD-401-3001
For Cluster Generation on cBot and Sequencing on the Genome Analyzer IIx	
TruSeq Paired-End Cluster Kit v2—cBot—GA (1 flow cell)	PE-300-2001
TruSeq Single-Read Cluster Kit v2—cBot—GA (1 flow cell)	GD-300-2001
For Cluster Generation on the Cluster Station and Sequencing on the Genome Analyzer IIx	
TruSeq Paired-End Cluster Kit v5—CS—GA (1 flow cell)	PE-203-5001
TruSeq Single-Read Cluster Kit v5—CS—GA (1 flow cell)	GD-203-5001

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

Pub. No. 970-2009-039 Current as of 27 April 2011

illumina®

Data Sheet: Sequencing

illumina®

TruSeq® Stranded mRNA and Total RNA Sample Preparation Kits

The clearest and most complete view of the transcriptome with a streamlined, cost efficient, and scalable solution for mRNA or whole-transcriptome analyses.

Highlights

- **Precise Measurement of Strand Orientation**
  Enables detection of antisense transcription, enhances transcript annotation, and increases alignment efficiency
- **Unparalleled Coverage Quality**
  High coverage uniformity enables most accurate and complete mapping of alternative transcripts and gene fusions
- **Configurations Compatible with Many Sample Types Including Low-Quality, FFPE, and Blood Samples**
  Leverage the power of RNA-Seq for previously inaccessible samples

Introduction

RNA sequencing (RNA-Seq) is a powerful method for discovering, profiling, and quantifying RNA transcripts. Using Illumina next generation sequencing technology, RNA-Seq does not require species- or transcript-specific probes, meaning the data are not biased by previous assumptions about the transcriptome. RNA-Seq enables hypothesis-free experimental designs of any species, including those with poor or missing genomic annotation. Beyond the measurement of gene expression changes, RNA-Seq can be used for discovery applications such as identifying alternative splicing events, gene fusions, allele-specific expression, and examining rare and novel transcripts.

As the complexities of gene regulation become better understood, a need for capturing additional data has emerged. Stranded information identifies from which of the two DNA strands a given RNA transcript was derived. This information provides increased confidence in transcript annotation, particularly for non-human samples. Identifying strand origin increases the percentage of alignable reads, reducing sequencing costs per sample. Maintaining strand orientation also allows identification of antisense expression, an important mediator of gene regulation[1]. The ability to capture the relative abundance of sense and antisense expression provides visibility to regulatory interactions that might otherwise be missed.

As the important biological roles of noncoding RNA continue to be recognized, whole-transcriptome analysis, or total RNA-Seq, provides a broader picture of expression dynamics. Total RNA-Seq enabled by ribosomal RNA (rRNA) reduction is compatible with formalin-fixed paraffin embedded (FFPE) samples, which contain potentially critical biological information. The family of TruSeq Stranded Total RNA sample preparation kits provides a unique combination of unmatched data quality for both mRNA and whole-transcriptome analyses, robust interrogation of both standard and low-quality samples and workflows compatible with a wide range of study designs (Figure 1).

Effective Ribosomal Reduction

TruSeq Stranded Total RNA kits couple proven ribosomal reduction and sample preparation chemistries into a single, streamlined workflow. Unlike polyA-based capture methods, Ribo-Zero kits remove ribosomal RNA (rRNA) using biotinylated probes that selectively bind rRNA species. The probe:rRNA hybrid is then captured by magnetic beads and removed, leaving the desired rRNA-depleted RNA in solution. This process minimizes ribosomal contamination and maximizes the percentage of uniquely mapped reads covering both mRNA and a broad range of non-coding RNA species of interest, including long intergenic noncoding RNA (lincRNA), small nuclear (snRNA), small nucleolar (snoRNA), and other RNA species[2].

High Quality Stranded Information

TruSeq Stranded RNA kits deliver unmatched data quality. The stranded measurement, or the percentage of uniquely mapped reads that return accurate strand origin information based on well-characterized universal human reference (UHR) RNA, is ≥ 99% using Stranded mRNA and ≥ 98% using Stranded Total RNA. This highly accurate information serves to increase the percentage of uniquely alignable reads in the assembly of poorly annotated transcriptomes and provides sensitivity to detect antisense expression. Consistent, precise measurement of RNA abundance is reflected by high reproducibility between technical replicates (Figure 2, R^2 = 0. 9873).

Figure 1: TruSeq Stranded RNA Sample Preparation Kits

The TruSeq Stranded mRNA and Total RNA Kits allow robust interrogation of both standard and low-quality samples, and include workflows compatible with a wide range of study designs.

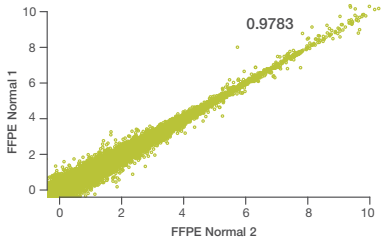
TruSeq Total RNA for Low-Quality Samples

TruSeq Total RNA enables robust and efficient interrogation of FFPE and other low-quality RNA samples. As shown in Figure 3, coverage across transcripts is high and even in both fresh-frozen (FF) and FFPE samples prepared with the TruSeq Stranded Total RNA kit. The optimized Ribo-Zero™ rRNA removal workflow provides a viable, highly scalable solution for efficient whole transcriptome analysis across samples that have been historically difficult to analyze.

RNA Analysis of Blood Samples

TruSeq Stranded Total RNA kits with Ribo-Zero Globin enable the efficient, robust interrogation of coding and noncoding RNA isolated from blood samples. A streamlined, automation-friendly workflow applies Ribo-Zero chemistry to simultaneously remove globin mRNA along with both cytoplasmic and mitochondrial rRNA in a single, rapid step (Table 1). In comparison to library preparation after ribosomal RNA reduction only, TruSeq Stranded Total RNA kits with Ribo-Zero Globin reduced globin mRNA levels generated from commercially obtained, blood-derived RNA from 28% to only 0.3% of aligned reads. These kits combine globin mRNA removal, rRNA removal, and library preparation to optimize sequencing output while reducing total assay time, eliminating the need for additional removal chemistry and reducing costs per sample.

Figure 2: Technical Replicates



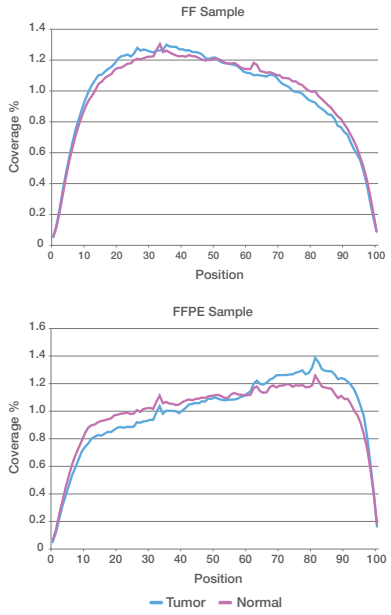
Technical replicates of FFPE tissue show high concordance, indicating robust sample prep performance. Axes are log2(FPKM). R^2 value is shown.

Differential Expression of Noncoding RNA

Maintaining strand information of RNA transcripts is important for many reasons. The example in Figure 4 shows a differentially-expressed transcript of the *ATP5H* gene in breast tumor and normal tissue prepared using the TruSeq RNA with Ribo-Zero compared to a standard polyA-based method. Both TruSeq Stranded Total RNA and polyA-prepared samples detect the differential expression of ATP5H between tumor and normal samples. However, using the Stranded Total RNA sample preparation kit, differential expression in reverse orientation at the position of pseudogene transcript AC087651.1 is also detected in the expected, opposite strand orientation.

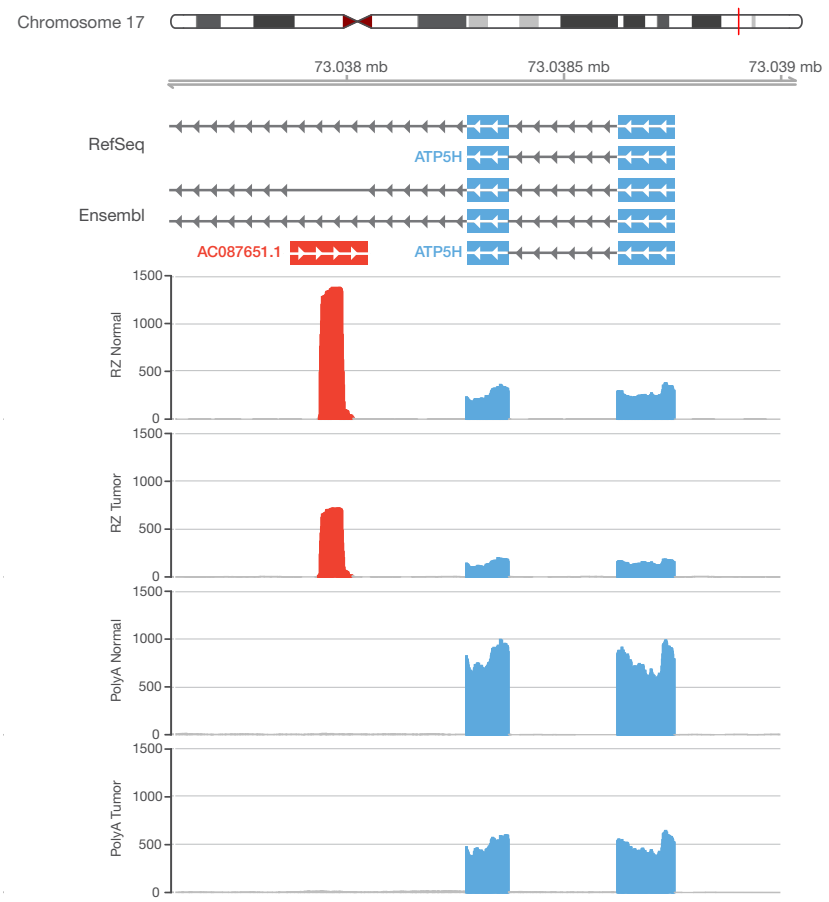
The example in Figure 5 shows that TruSeq Stranded Total RNA enables reliable detection of differential expression across multiple forms of ncRNA, including lincRNA, snRNA, snoRNA, and other RNA species.

Figure 3: Even Coverage Across Transcripts



TruSeq Stranded Total RNA gives excellent coverage across the top 1,000 expressed transcripts in both fresh-frozen (FF, top) and FFPE (bottom) tumor and matched normal breast tissue, with > 98% aligned stranded reads. X-axis: position along transcript, Y-axis = percent coverage of combined reads.

Figure 4: Differential Expression of ncRNA Transcripts



ATP5H expression from chromosome 17 is differentially expressed in breast tumor vs. normal tissue. Using two different sample preparation methods (RZ; Ribo-Zero for total RNA or PolyA-based mRNA) shows differential expression in tumor vs. normal tissues in both preps (Blue). However, only Total RNA with Ribo-Zero reveals differential expression at the locus of a pseudogene (Red, AC087651.1), for which reads are detected in the opposite orientation, as expected. This stranded information would have been lost in a standard mRNA prep.

Table 1: Targeted RNA Species

Kit Name	Cytoplasmic rRNA	Mitochondrial rRNA	Globin mRNA
TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero Human/Mouse/Rat	Targeted	Not targeted	Not targeted
TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero Gold	Targeted	Targeted	Not targeted
TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero Globin	Targeted	Targeted	Targeted

Several TruSeq Stranded Total RNA with Ribo-Zero kit configurations are available to suit a range of study designs, providing highly efficient removal of cytoplasmic rRNA, cytoplasmic and mitochondrial rRNA, or both forms of rRNA in addition to globin mRNA.

Flexible Workflow Configurations

The TruSeq Stranded mRNA and Total RNA kits offer solutions optimized for your individual experimental needs. Each kit includes two workflows: the high throughput protocol is ideally suited for projects with ≥ 48 samples, and the low throughput protocol is best suited for projects with ≤ 48 samples. Stranded Total RNA configurations are available for targeting the removal of either cytoplasmic rRNA only, or both cytoplasmic plus mitochondrial rRNA (Table 2). In a comparison using Universal Human Reference RNA, TruSeq Stranded Total RNA kits with Ribo-Zero Human/Mouse/Rat and Gold both reduced cytoplasmic rRNA to < 2% of aligned reads, whereas those with Ribo-Zero Gold additionally reduced mitochondrial rRNA from 7% to only 0.02% of aligned reads.

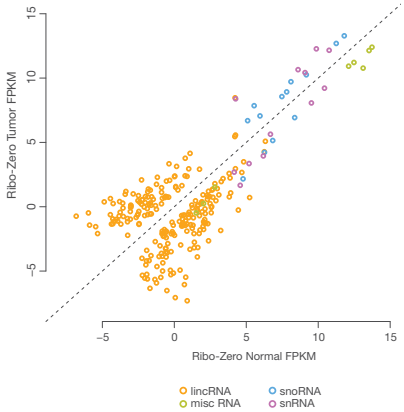
Conclusion

TruSeq Stranded mRNA sample prep kits provide the clearest, most complete view of the transcriptome, providing precise measurement of strand orientation, uniform coverage, and high-confidence discovery of features such as alternative transcripts, gene fusions, and allele-specific expression. TruSeq Stranded Total RNA kits couple all of the benefits of TruSeq RNA preparation kits with Ribo-Zero ribosomal reduction chemistry, providing a robust and highly scalable end-to-end solution for whole-transcriptome analysis compatible with a wide range of samples, including non-human and FFPE.

References

1. Nagai K, Kohno K, Chiba M, Pak S, Murata S, et al. (2012) Differential expression profiles of sense and antisense transcripts between HCV-associated hepatocellular carcinoma and corresponding non-cancerous liver tissue. Int J Oncol 40(6):1813–20.
2. Ribo-Zero Gold Kit: Improved RNA-Seq results after removal of cytoplasmic and mitochondrial ribosomal RNA. Nature Methods Application Note, 2011.

Figure 5: Detection of ncRNA Expression



With TruSeq Stranded Total RNA sample preparation, differential expression across a range of non-coding RNA species, including long intergenic noncoding RNA (lincRNA), small nuclear (snRNA) and small nucleolar (snoRNA) and other species (misc RNA) can be detected between tumor and normal tissues (four replicates per sample, false discovery rate (FDR) = 0.05).

Data Sheet: Sequencing

illumina®

TruSeq® Targeted RNA Expression

Highly customizable and affordable mid-plex gene expression analysis for the MiSeq® system.

Highlights

- **Content and flexibility with fixed and customizable panels**
  Choose validated pathway, cell, or disease-specific fixed panels, or add customized content
- **Mid-plex gene expression at a complexity and scale not previously possible**
  Examine 1,000 targets per sample, 384 samples per run
- **Fast and simple workflow**
  Go from RNA to data in less than two days

Introduction

TruSeq Targeted RNA Expression leverages proven MiSeq sequencing technology to deliver an accurate and powerful method for validating gene expression arrays and RNA-Seq studies. TruSeq Targeted RNA Expression (Figure 1) enables efficient, quantitative multiplexed gene expression profiling for 12-1,000 targets per sample and up to 384 samples in a single MiSeq run. Requiring just 50 ng or less of starting RNA, TruSeq Targeted RNA Expression is amenable to a wide range of samples. Choose from over 400,000 pre-designed assays to create a custom panel targeting genes, exons, splice junctions, cSNPs and fusions. Fixed panels offer a wide variety of biological pathways and disease-specific markers, or combine fixed and custom content for the ultimate in flexibility. TruSeq Targeted RNA Expression offers a fully integrated solution, including convenient online assay design and ordering, a streamlined workflow, and automated, on-instrument data analysis.

Choose Fixed Panels for Focused Studies

For pathway- or disease-focused expression or profiling studies, TruSeq Targeted RNA Expression fixed panels offer ready-to-use assays designed for commonly studied genes (Table 1). Validated, fixed content panels are ideal for profiling many samples or screening cell types quickly and economically, and providing base content that can be expanded upon with custom content as needed.

Increase Your Flexibility with Custom Content

TruSeq Targeted RNA Expression assays are pre-designed assays targeting exon junctions and non-junction sites, as well as target SNPs within coding regions. Choose validated assays in DesignStudio™, a free, user-friendly tool accessed through your MyIllumina account[1].

Figure 1: TruSeq Targeted RNA Expression

TruSeq Targeted RNA Expression delivers fixed or customizable affordable mid-plex gene expression that takes full advantage of the throughput and flexibility of the MiSeq® system.

Create fully custom panels of 12–1,000 assays, or add specific genes or regions to one of the fixed panels, or to a previously ordered custom panel. Simply select the assays you need and add them to your order, with no design time.

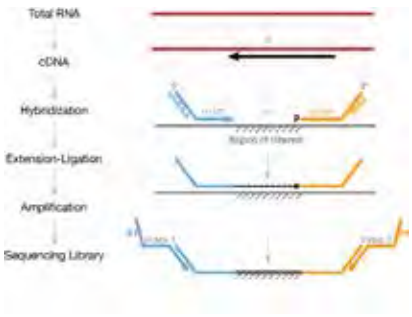
Streamlined, Targeted Assay Workflow

TruSeq Targeted RNA Expression for custom or fixed designs features a simple method for generating indexed, sequence-ready libraries from RNA regions of interest (Figure 2). Starting with as little as 50 ng of total RNA, the small amplicon size allows successful target detection, even on poor quality samples. All targets are amplified in a single reaction, minimizing potential bias and workflow steps compared to methods such as qPCR. From sample to data analysis, the entire process takes less than two days.

Table 1: TruSeq RNA Expression Fixed Panels

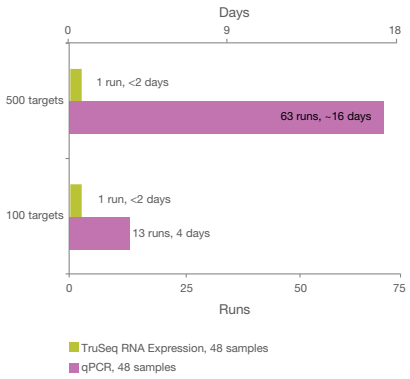
Apoptosis	Hedgehog Pathway	TP53 Pathway
Cardiotoxicity	Neurodegeneration	Wnt Pathway
Cell Cycle	NFκB Pathway	
Cytochrome P450	Stem Cell	

Figure 2: TruSeq Targeted RNA Expression Workflow



The TruSeq Targeted RNA Expression assay chemistry begins with reverse transcribing cDNA from purified total RNA. Two custom-designed oligonucleotide probes with adapter sequences hybridize up and downstream of the region of interest. An extension-ligation reaction, followed by amplification creates a new template strand. Templates are then PCR amplified to add indices, creating sequence-ready libraries.

Figure 3: TruSeq Targeted RNA vs qPCR Workflow



With TruSeq RNA Expression, run 500 targets on 48 samples in one run in less than two days, compared to 63 runs in ~16 days with qPCR methods.

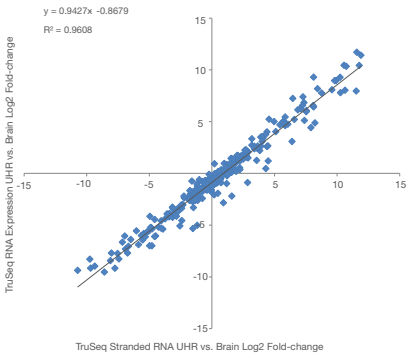
Multiplexing at a Scale not Previously Possible

With TruSeq Targeted RNA Expression, you can run up to 384 dual-index combinations to efficiently multiplex samples within a single MiSeq run. With 25 million reads, the MiSeq system is capable of generating 25,000 datapoints per run (at an average of 1,000 reads per target), equivalent to 65 384-well plates. Compared to qPCR, the number of runs and amount of processing time is significantly decreased (Figure 3). For more information about read budget, normalization, and getting the best results from your TruSeq Targeted RNA Expression assays, refer to the technical note[2].

Accurate Confirmation Using TruSeq Targeted RNA Expression

TruSeq Targeted RNA Expression was compared against the gold standard RNA-Seq for fold-change in an experimental target set. As shown in Figure 4, fold-change expression in 281 targets between Universal Human Reference (UHR) RNA and total brain mRNA was measured using TruSeq Targeted RNA Expression (X-axis) and TruSeq Stranded RNA-Seq (Y-axis). Data show excellent correlation, demonstrating that TruSeq Targeted RNA Expression provides accurate validation. The assay is also highly reproducible, even over a large dynamic range (Figure 5).

Figure 4: Fold-Change Correlation between RNA-Seq and TruSeq Targeted RNA Expression



Comparison of fold change expression between Universal Human Reference (UHR) and brain mRNAs for 281 targets, using TruSeq Stranded RNA-Seq (X-axis) and TruSeq RNA Expression (Y-axis).

Figure 5: Visualization of TruSeq Targeted RNA Expression Data using MiSeq Reporter

Data visualization with MiSeq Reporter allows easy comparison of data sets.

Product Specifications

Specification	Value
Database content	> 400,000 designs (mouse, human, rat)
Target types	Gene, transcript, exon, splice junction, cSNP, fusion
Dynamic range	5 orders of magnitude
Time to answer	1.5 days
Hands-on time	4 hours
RNA quality	> 200 bp unfixed or FFPE

Simple Data Analysis

After a sequencing run on the MiSeq system, data are automatically aligned and can be viewed using the MiSeq Reporter. As shown in Figure 5, pairwise comparisons for relative expression between samples or groups of samples is simple and intuitive. Customizable significance thresholds allow you to quickly identify differentially expressed targets. The TruSeq Targeted RNA Expression user experience is customized and streamlined, and keeps project data highly accessible.

Summary

Designed for the MiSeq system, TruSeq Targeted RNA Expression provides rapid and economical RNA profiling and validation for your gene expression studies. Go from sample to answer in less than two days with a simple, streamlined workflow and automated data visualization. Choose validated, pre-designed panels or add custom content to your existing assays for the ultimate flexibility to evolve your research.

References

1. https://icom.illumina.com/
2. Considerations for Designing a Successful TruSeq Targeted RNA Expression Experiment Technical Note, 2013.

Ordering Information

Product Name	Number of Samples	Catalog No.
TruSeq Targeted RNA Expression Custom Components		
TruSeq Targeted RNA Custom Kit	48	RT-101-1001
	96	RT-102-1001
TruSeq Targeted RNA Supplemental Content	48	RT-801-1001
	96	RT-802-1001
TruSeq Targeted RNA Expression Fixed Panels		
TruSeq Targeted RNA Apoptosis Panel Kit	48	RT-201-1010
	96	RT-202-1010
TruSeq Targeted RNA Cardiotoxicity Panel Kit	48	RT-201-1009
	96	RT-202-1009
TruSeq Targeted RNA Cell Cycle Panel Kit	48	RT-201-1003
	96	RT-202-1003
TruSeq Targeted RNA Cytochrome p450 Panel Kit	48	RT-201-1006
	96	RT-202-1006
TruSeq Targeted RNA Hedgehog Panel Kit	48	RT-201-1002
	96	RT-202-1002
TruSeq Targeted RNA Neurodegeneration Panel Kit	48	RT-201-1001
	96	RT-202-1001
TruSeq Targeted RNA NFκB Panel Kit	48	RT-201-1008
	96	RT-202-1008
TruSeq Targeted RNA Stem Cell Panel Kit	48	RT-201-1005
	96	RT-202-1005
TruSeq Targeted RNA TP53 Pathway Panel Kit	48	RT-201-1007
	96	RT-202-1007
TruSeq Targeted RNA Wnt Pathway Panel Kit	48	RT-201-1004
	96	RT-202-1004
TruSeq Targeted RNA Expression Index Kits		
TruSeq Targeted RNA Index Kit	48	RT-401-1001
TruSeq Targeted RNA Index Kit A	96	RT-402-1001
TruSeq Targeted RNA Index Kit B	96	RT-402-1002
TruSeq Targeted RNA Index Kit C	96	RT-402-1003
TruSeq Targeted RNA Index Kit D	96	RT-402-1004

Illumina • 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. 470-2012-004 Current as of 14 August 2013

illumina®

epicentre®
an Illumina® company

RNA-Seq of Blood

ScriptSeq™ Complete Gold (Blood)

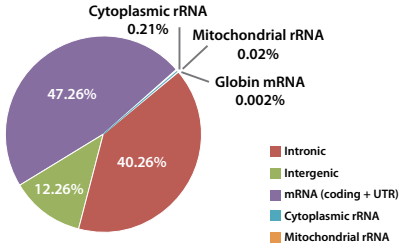
- Removes globin mRNA and ribosomal RNA
- Creates an Illumina® sequencing library
- The data contains high amounts of coding and non-coding information

- Find more genes
- Find more coding and non-coding RNA's
- Good for small samples
- All phases of research

Blood is an important sample for research into 6,000 rare diseases and 12,000 disease groups. The data from samples treated with ScriptSeq Complete Gold (Blood) is focused on valuable RNA. Finding new genes, splice variants and isoforms is important to disease and health research.

ScriptSeq Complete (Blood) offers the most informative sequencing results by removing unwanted globin mRNA and ribosomal RNA prior to sequencing.

Figure 1. RNA-Seq libraries contain coding and non-coding RNA.



Find more coding and non-coding RNA

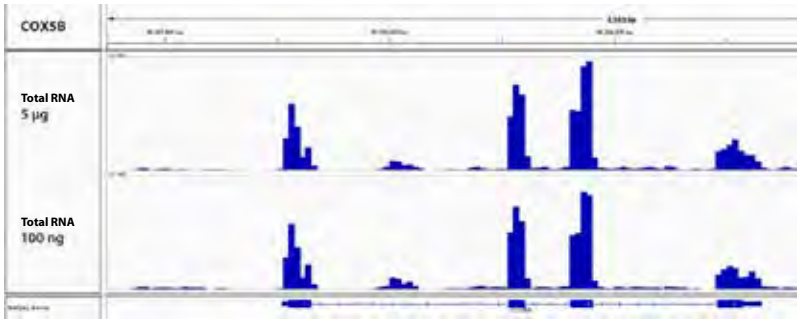
ScriptSeq Complete Gold (Blood) libraries were prepared from 5 µg of RNA isolated from human whole blood and sequenced on an Illumina® sequencer. Greater than 98% of all reads contain useful information.

RNA-Seq data is very useful to study disease (or health). Figure 1 shows an example in which 47% of the sequencing reads contain coding RNA and 52% contain non-coding RNA.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 001

Figure 2. Gene coverage from large (5 µg) or small (100 ng) of RNA.



Available for all sample sizes

ScriptSeq Complete Gold (Blood) is available for 100 ng + of total RNA. Results from small amounts of total RNA are very similar to results from high amounts of total RNA. Figure 2 shows coverage of the COX5B gene when either a small amount (100 ng) of total RNA or large amount (5 µg) of total RNA was treated with ScriptSeq Complete Gold (Blood).

Strong gene coverage

In figure 2, the height of the blue bars show how many reads align to that sequence. Taller bars show more reads and deeper (better) coverage. Coding (thick blue bars) regions in both the small and large input ranges is similar.

Success begins with purification

MasterPure RNA purification kit

Purification is an important step to prepare your sample. MasterPure safely removes unwanted material to give you pure, intact total RNA.

MasterPure offers unique benefits:

- Keep RNA intact (does not degrade RNA)
- Retain RNA diversity (including small RNA)
- Maximize genes discovered
- Available for all sample sizes

Cat. #	Quantity
MasterPure™ RNA Purification Kit (for isolating RNA only)	
MCR85102	100 Purifications

Cat. #	Quantity
ScriptSeq™ Complete Gold Kit (Blood)—Low Input	
SCL24GBL	24 Reactions
SCL6GBL	6 Reactions
For 100 ng – 1 µg total blood RNA.	
ScriptSeq™ Complete Gold Kit (Blood)	
BGGB1306	6 Reactions
BGGB1324	24 Reactions
For 1 µg – 5 µg total blood RNA.	
FailSafe™ PCR Enzyme Mix	
FSE51100	100 Units

Patents: www.illumina.com/patents

epicentre®
an Illumina® company

RNA-Seq of Yeast

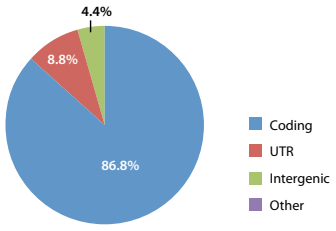
ScriptSeq™ Complete Gold Kit (Yeast)

- Removes ribosomal RNA with Ribo-Zero™
- Creates an Illumina® sequencing library with ScriptSeq v2
- Results contain coding and non-coding RNA

- One day method
- Find more genes
- Good for small samples

The yeast transcriptome is more complex than previously thought. RNA-Seq of yeast is a valuable approach for mapping the transcriptome and characterizing novel and low abundance transcripts. The ScriptSeq Complete Gold Kit (Yeast) offers the most informative sequencing results by removing unwanted ribosomal RNA prior to sequencing.

Figure 1. RNA-Seq libraries contain coding and non-coding RNA.



Library composition of ScriptSeq™ Complete Gold Kit (Yeast) samples. ScriptSeq libraries were constructed from 1 µg of *S. cerevisiae* total RNA samples and sequenced on an Illumina® MiSeq™.

Find more coding and non-coding RNA

ScriptSeq Complete Gold Kit (Yeast):

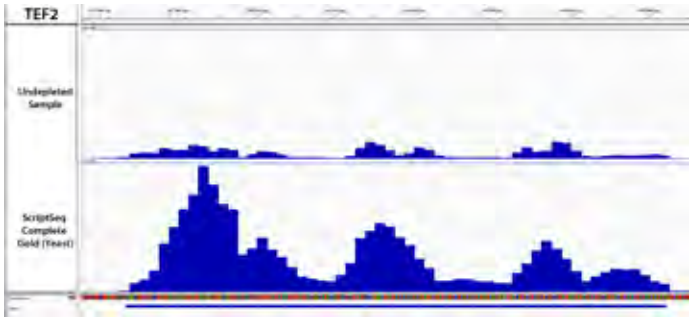
- Find more coding RNA
- Removes ribosomal RNA
- Creates Illumina® sequencing libraries
- Data contains high amounts of coding information

RNA-Seq data is very useful to study yeast gene expression. Figure 1 shows an example in which 95.6 % of the sequencing reads contain coding RNA and 4.4 % contain non-coding RNA.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 002

Figure 2. Enhanced coverage with ScriptSeq™ Complete Gold Kit (Yeast).



Enhanced coverage with ScriptSeq Complete Gold Kit (Yeast)

ScriptSeq Complete Gold (Yeast) contains Ribo-Zero Gold (Yeast) for depletion of yeast rRNA. Gene coverage of the TEF2 gene (Figure 2) shows that rRNA depletion reveals more reads. In the figure, the height of the blue bars shows how many reads align to that sequence. Taller bars show more reads and deeper (better) coverage.

ScriptSeq Yeast is a powerful tool to study:

- Transcriptome mapping
- Gene structure
- Characterization of novel and low abundance transcripts

Cat. #	Quantity
Ribo-Zero™ Magnetic Gold Kit (Yeast)	
Suitable for 1-5 µg of total RNA.	
MRZY1306	6 Reactions
MRZY1324	24 Reactions
ScriptSeq™ Complete Gold Kit (Yeast)	
Includes Ribo-Zero Gold (Yeast). Suitable for 1-5 µg of total RNA.	
BGY1306	6 Reactions
BGY1324	24 Reactions
ScriptSeq™ Complete Gold Kit (Yeast)- Low Input	
Includes Ribo-Zero Gold (Yeast). Suitable for 100 ng - 1 µg of total RNA.	
SCGL6Y	6 Reactions
SCGL6Y	24 Reactions
FailSafe™ PCR Enzyme Mix	
FSE51100	100 Units

The FailSafe PCR Enzyme Mix is required for ScriptSeq Complete Gold (Yeast) RNA-Seq library preparation.

Success begins with purification

MasterPure™ RNA Purification Kit

Purification is the first critical step to prepare samples for sequencing. MasterPure produces sequencer-ready RNA safely and easily.

MasterPure offers unique benefits:

- Keep RNA intact (does not degrade RNA)
- Retain RNA diversity (including small RNA)
- Maximize genes discovered
- Available for all sample sizes

Cat. #	Quantity
MasterPure™ RNA Purification Kit (for isolating RNA only)	
MCR85102	100 Purifications

epicentre®
an Illumina® company

Ribosome Profiling

ARTseq™ Ribosome Profiling Kits

ARTseq (**A**ctive m**R**NA **T**ranslation) Ribosome profiling is a powerful technique to study translation.

- Sequence ribosome protected mRNA
- Rapid, scalable spin-column method
- No ultracentrifuge required!
- Compatible with yeast and mammalian samples

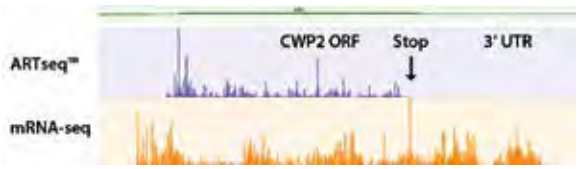
- Predict protein abundance
- Investigate translational control
- Measure gene expression

Sequencing actively translated transcripts

Sequence mRNA fragments undergoing translation by ribosomes. These mRNA fragments are called "footprinted" or ribosome protected mRNA fragments.

ARTseq sequences ribosome-protected mRNA fragments to provide a "snapshot" of the active ribosomes in a cell. You can identify proteins being actively translated from samples prepared with ARTseq. Samples collected at different times often show changes in translation. Samples treated with different drugs often show different translation patterns.

Figure 1. Identify actively translated RNA's with ARTseq.



Sequence only the protein coding regions

Samples prepared with ARTseq are enriched for ORF and devoid of UTR sequences. The start and stop codons are easily seen. Sequences are focused on protein coding regions.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 006

Infinium HumanMethylation450 BeadChip

Data Sheet: Epigenetics

illumina®

Infinium® HumanMethylation450 BeadChip

The ideal solution for affordable, large sample–size genome-wide DNA methylation studies.

HumanMethylation450 BeadChip Highlights

- **Unique Combination of Genome-Wide Coverage, High-Throughput, and Low Cost**
  Over 450,000 methylation sites per sample at single-nucleotide resolution
- **Unrivaled Assay Reproducibility**
  > 98% reproducibility for technical replicates
- **Simple Workflow**
  PCR-free protocol with the powerful Infinium HD Assay
- **Compatibile with FFPE Samples**
  Protocol available for methylation studies on FFPE samples

Introduction

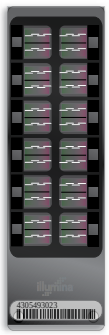
DNA methylation plays an important and dynamic role in regulating gene expression. It allows cells to become specialized and stably maintain those unique characteristics throughout the life of the organism, suppresses the deleterious expression of viral genes and other non-host DNA elements, and provides a mechanism for response to environmental stimuli. Aberrant DNA methylation (hyper- or hypomethylation) and its impact on gene expression have been implicated in many disease processes, including cancer[1].

To enable cost-effective DNA methylation analysis for a variety of applications, Illumina offers a robust methylation profiling platform consisting of proven chemistries and the iScan and HiScan®SQ systems. The HumanMethylation450 BeadChip (Figure 1) offers a unique combination of comprehensive, expert-selected coverage and high throughput at a low price, making it ideal for screening large sample populations such as those used in genome-wide association study (GWAS) cohorts. By providing quantitative methylation measurement at the single-CpG–site level for normal and formalin-fixed parafin-embedded (FFPE) samples, this assay offers powerful resolution for understanding epigenetic changes.

Comprehensive Genome-Wide Coverage

The Infinium HumanMethylation450 BeadChip provides unparalleled, genome-wide coverage featuring comprehensive gene region and CpG island coverage, plus additional high-value content selected with the guidance of methylation experts. Infinium HD technology enables content selection independent of bias-associated limitations often associated with methylated DNA capture methods. As a result, 99% of RefSeq genes are covered, including those in regions of low CpG island density and at risk for being missed by commonly used capture methods.

Figure 1: Infinium HumanMethylation450 BeadChip



The Infinium HumanMethylation450 BeadChip features more than 450,000 methylation sites, within and outside of CpG islands.

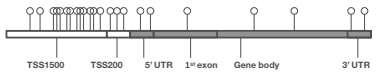
Importantly, coverage was targeted across gene regions with sites in the promoter region, 5'UTR, first exon, gene body, and 3'UTR in order to provide the broadest, most comprehensive view of methylation state possible (Figure 2). This multiple-site approach was extended to CpG islands/CpG island regions for which 96% of islands were covered overall, with multiple sites within islands and island shores, as well as those regions flanking island shores (island shelves). Beyond gene and CpG island regions, multiple additional content categories requested by methylation experts were also included:

- CpG sites outside of CpG islands
- Non-CpG methylated sites identified in human stem cells
- Differentially methylated sites identified in tumor versus normal (multiple forms of cancer) and across several tissue types
- FANTOM 4 promoters
- DNase hypersensitive sites
- miRNA promoter regions
- ~ 90% of content contained on the Illumina HumanMethylation27 BeadChip

Streamlined Workflow

The HumanMethylation450 BeadChip follows a user-friendly, streamlined workflow that does not require PCR. Its low sample input requirement (as low as 500 ng), enables analysis of valuable samples

Figure 2: HumanMethylation450 BeadChip Provides Coverage Throughout Gene Regions



Feature Type	Genes Mapped	Percent Genes Covered	Number of Loci on Array
NM_TSS200	14895	0.79	2.56
NM_TS1500	17820	0.94	3.41
NM_5'UTR	13865	0.78	3.34
NM_1stExon	15127	0.80	1.62
NM_3'UTR	13042	0.72	1.02
NM_GeneBody	17071	0.97	8.97
NR_TSS200	1967	0.65	1.84
NR_TSS1500	2672	0.88	2.92
NR_GeneBody	2345	0.77	5.34

Feature Type	Islands Mapped	Percent Islands Covered	Average Number of Loci on Array
Island	26153	0.94	5.08
N_Shore	25770	0.93	2.74
S_Shore	25614	0.92	2.66
N_Shelf	23896	0.86	1.97
S_Shelf	23968	0.86	1.94

The HumanMethylation450 BeadChip offers broad coverage across gene regions, as well as CpG islands/CPG island regions, shelves, and shores for the most comprehensive view of methylation state.

derived from limited DNA sources. HumanMethylation450 BeadChip kits contain all required reagents for performing methylation analyses (except for the bisulfite conversion kit, which is available separately).

Data Integration

Of all the genes represented on the HumanMethylation450 BeadChip, more than 20,000 are also present on the HumanHT-12 v4 Expression BeadChip[2], permitting combined analysis of global methylation status and gene expression levels. In addition, investigators may integrate methylation data with genotyping data from GWAS studies to better understand the interplay between genotype and methylation state in driving phenotypes of interest.

High-Quality Data

The HumanMethylation450 BeadChip applies both Infinium I and II assay chemistry technologies (Figure 3) to enhance the depth of coverage for methylation analysis. The addition of the Infinium II design allows use of degenerate oligonucleotide probes for a single bead type, enabling each of up to three underlying CpG sites to be either methylated or unmethylated with no impact on the result for the queried site.

Illumina scientists rigorously test every product to ensure strong and reproducible performance, enabling researchers to achieve industry-leading data quality.

Precision and Accuracy

Reproducibility has been determined based on the correlation of results generated from technical replicates. The HumanMethylation450 BeadChip showed strong correlation between replicates (r>0.98), as well as with the HumanMethylation27 BeadChip and whole-genome bisulfite sequencing (Figure 4).

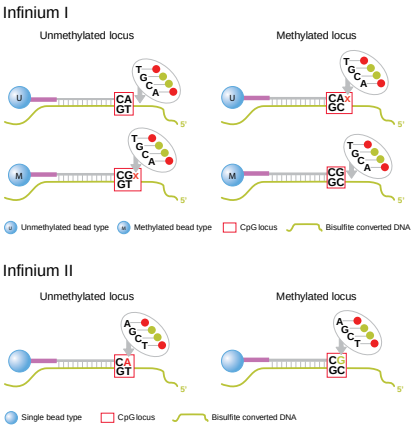
Sensitivity

By comparing the results of replicate experiments (duplicates of eight biological samples), Illumina scientists have shown that the HumanMethylation450 BeadChip reliably detects a delta-beta value of 0.2 with a lower than 1% false positive rate.

Internal Quality Controls

Infinium HD–based assays possess several sample-dependent and sample-independent controls so researchers have confidence in producing the highest quality data. The HumanMethylation450 BeadChip includes 600 negative controls, which are particularly important in methylation analysis assays since sequence complexity is decreased after bisulfite conversion. The GenomeStudio® Methylation Module Software has an integrated Controls Dashboard where the performance of all controls can be easily monitored.

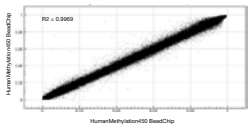
Figure 3: Broader Coverage Using Infinium I and II Assay Designs



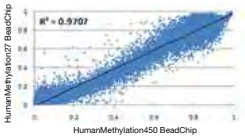
The HumanMethylation450 BeadChip employs both Infinium I and Infinium II assays, enhancing its breadth of coverage. Infinium I assay design employs two bead types per CpG locus, one each for the methylated and unmethylated states. The Infinium II design uses one bead type, with the methylated state determined at the single base extension step after hybridization.

Figure 4: High Assay Reproducibility

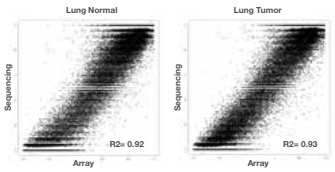
A: HumanMethylation450 Replicate Correlation



B: HumanMethylation27 vs. HumanMethylation450 Correlation

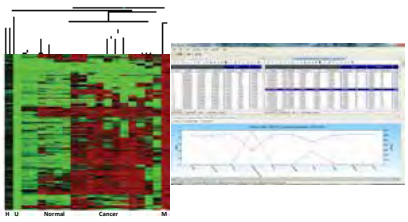


C. HumanMethylation450 vs. Whole-Genome Bisulfite Sequencing



Using the HumanMethylation450 BeadChip, users can be confident of obtaining consistent, robust data. Representative plots from internal testing show strong replicate correlation (A), as well as strong correlation with the HumanMethylation27 BeadChip (B) and whole-genome bisulfite sequencing (C).

Figure 5: Integrated Data Analysis with Illumina GenomeStudio Software



GenomeStudio software supports DNA methylation analysis on any platform. Data are displayed in intuitive graphics. Gene expression data can be easily integrated with methylation projects (plotted on right).

Integrated Analysis Software

HumanMethylation450 BeadChip data analysis is supported by the powerful and intuitive GenomeStudio Methylation Module, enabling researchers to effortlessly perform differential methylation analysis (Figure 5). The GenomeStudio software features advanced visualization tools that enable researchers to view vast amounts of data in a single graph, such as heat maps, scatter plots, and line plots. These tools and the GenomeStudio Genome Browser display valuable information such as chromosomal coordinates, percent GC, location in a CpG Island, and methylation β values.

Data generated by the Infinium HD methylation assay are easily compatible with data from other Illumina applications, including gene expression profiling. This enables researchers to perform cross-application analysis such as the integration of gene expression data with HumanMethylation450 BeadChip methylation data.

Methylation Studies with FFPE Samples

Researchers can perform methylation studies on FFPE samples by using a special, modified version of the Infinium HumanMethylation450 BeadChip protocol[3] that leverages the easy-to-use Infinium FFPE DNA Restoration Solution[4] to produce robust, highly reproducible results (Table 1). The FFPE DNA Restoration Solution includes the Illumina FFPE QC and the Infinium HD FFPE DNA Restore Kits. Please note that while the FFPE DNA Restoration Solution and HumanMethylation450 BeadChip kits are the same for normal and FFPE samples, investigators running FFPE samples should only follow the workflow described in the Infinium HD FFPE Methylation Assay protocol (manual or automated)[5,6], as it includes important changes to the standard protocols for each kit.

Table 1: Comparative Infinium HumanMethylation450 Data Quality Metrics—Standard vs. FFPE

HumanMethylation450 BeadChip	Standard Protocol	FFPE Protocol
Reproducibility (Technical replicates)	$r^2 \geq 98\%$	$r^2 \geq 98\%$
Number of sites detected*	≥99%	≥ 95%

*Based on non-cancer samples, recommended sample input amounts of high-quality DNA as confirmed by PicoGreen and following all other Illumina recommendations as per respective User Guides.

Ordering Information

Catalog No.	Product	Description
WG-314-1003	Infinium HumanMethylation450 BeadChip Kit (24 samples)	Each package contains two BeadChips and reagents for analyzing DNA methylation in 24 human DNA samples.
WG-314-1001	Infinium HumanMethylation450 BeadChip Kit (48 samples)	Each package contains four BeadChips and reagents for analyzing DNA methylation in 48 human DNA samples.
WG-314-1002	Infinium HumanMethylation450 BeadChip Kit (96 samples)	Each package contains eight BeadChips and reagents for analyzing DNA methylation in 96 human DNA samples.

Each HumanMethylation450 BeadChip can process 12 samples in parallel and assay >450,000 methylation sites per sample.

Summary

The HumanMethylation450 BeadChip's unique combination of comprehensive, expert-selected coverage, high sample throughput capacity, and affordable price makes it an ideal solution for large sample–size, genome-wide DNA methylation studies.

References

1. Portela A, Esteller M (2010) Epigenetic modifications and human disease. Nat Biotechnology 28: 1057–1068.
2. http://www.illumina.com/products/humanht_12_expression_beadchip_kits_v4.ilmn
3. Infinium HD FFPE DNA Restoration Protocol
4. http://www.illumina.com/products/infinium_ffpe_dna_restoration_solution.ilmn
5. Infinium HD FFPE Methylation Assay, Manual Protocol
6. Infinium HD FFPE Methylation Assay, Automated Protocol
7. Illumina FFPE QC Assay Protocol

Illumina • +1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

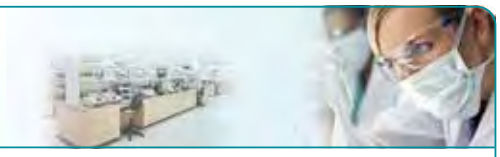
FOR RESEARCH USE ONLY

Pub. No. 270-2010-001 Current as of 09 March 2012

illumina®

FailSafe™ PCR System

epicentre®
an Illumina® company



PCR Optimization

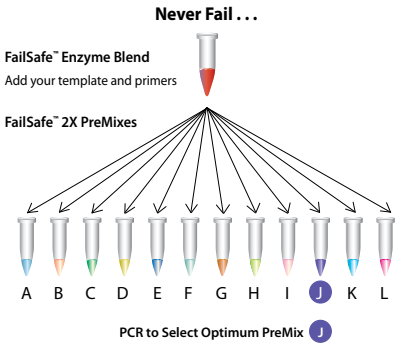
FailSafe™ PCR System

- PCR of difficult or high GC templates
- PCR Amplifications up to 20 kb
- Works the first time, every time
- 3-fold lower error rate than *Taq* DNA Polymerase

The FailSafe PCR System uses the patented Epicentre PCR enhancement technology to allow PCR reactions to work the first time and every time. Twelve buffer options run at the same time, allowing quick and easy optimization of your PCR.

FailSafe has enabled many difficult samples to be used successfully in PCR and published. FailSafe will ensure your PCR is successful.

Figure 1. Ensure successful PCR with FailSafe.



Works the first time, every time

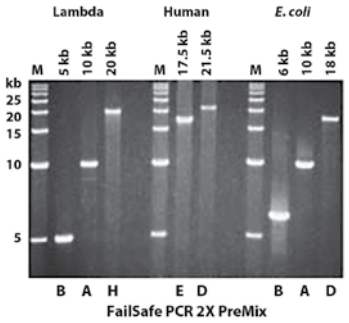
Optimizing PCR is easy with FailSafe. Create a master mix of FailSafe Enzyme blend, template DNA and primers. Add the FailSafe PCR PreMix Selection Kit buffers to test which buffer is optimal for your reaction.

Workflow

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089

PDS 008

Figure 2. The FailSafe™ PCR System will work with nearly all DNA – animal, bacterial, plant or viral.



FailSafe PCR will amplify DNA from a range of different sequences and sequence sizes. PCR products shown are up to 20 kb for lambda DNA, up to 21.5 kb for human DNA, and up to 18 kb for *E. coli* DNA.

Amplify any sample

FailSafe will amplify DNA sequences from almost any source, up to 20 kb in length in a single round.

FailSafe PCR products can be used in many applications, including TA cloning and blunt-end cloning.

FailSafe components (Polymerase and Premixes) are available in the FailSafe Premix Choice kits and as single products.

Cat. #	Quantity
FailSafe PCR PreMix Selection Kit	
FS99060	(Contains all 12 Premixes and FailSafe PCR Polymerase) – sufficient reagent for 48 reactions (four full template and primer optimizations)
FailSafe PCR System	
FS99100	(100 units of FailSafe Polymerase and 1 PreMix of choice)
FS99250	(250 units of Failsafe Polymerase and two Premixes of choice)
FS9901K	(1000 U of FailSafe Polymerase and eight PreMixes of choice)
FailSafe PCR Polymerase	
FSE51100	(100 U)
FailSafe PCR Polymerase	
FSE5101K	(1000 U)
FailSafe PCR Premixes	
FSP995A-L	(A through L), 2.5 ml (100 reactions)

FailSafe™ PCR Premix Selection Kits: Purchase of this product includes an immunity from suit under patents specified in the product insert to use only the amount purchased for the purchaser's own internal research. No other patent rights (such as 5' Nuclease Process patent rights) are conveyed expressly, by implication, or by estoppel. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 944.

Success begins with MasterPure™ DNA Purification Kit

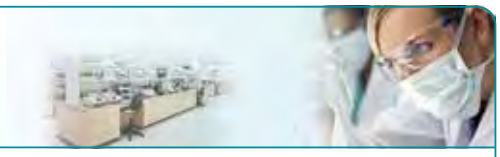
MasterPure produces DNA safely and easily to be used in many applications.

MasterPure offers unique benefits:

- Keep DNA intact (does not degrade DNA)
- Scalable reaction sizes
- Available for multiple sample types

Cat. #	Quantity
MasterPure™ Complete DNA and RNA Purification Kit	
MC85200	200 Purifications
MC89010	10 Purifications

epicentre®
an Illumina® company



What Will You Create Today?

Innovative Enzyme Solutions

- Stringent QC
- No affinity tags
- ISO 13485 compliant by end of 2013

Example applications:

- Reverse transcriptases
- RNA polymerases
- RNase-free DNase
- Much more...

Epicentre develops and manufactures the highest purity enzymes for life science research. Epicentre specializes in unique and bulk enzyme projects to meet your specific needs. Since opening in 1987, Epicentre has a proven track record in manufacturing high purity enzymes for life sciences in our state-of-the-art facility in Madison, Wisconsin, USA.

Standard enzymes for molecular biology research are available. Unique, hard-to-find enzymes are available to your specifications.

OEM opportunities available

- Custom manufacturing for alternate size requirements or specific concentrations
- In-house technical expertise with over 25 years of experience
- Competitive pricing to meet your budget constraints
- Flexible, quick turnaround time for OEM needs

Bulk availability

- Standard or custom offerings
- Flexible (custom) concentrations and package sizes
- Bulk capabilities

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089 PDS 005

Browse the possibilities…

DNA Polymerases

Klenow DNA Polymerase

Exo-Minus Klenow DNA Polymerase (D355A, E357A)

RepliPHI™ Phi29 DNA Polymerase

Terminal deoxynucleotidyl Transferase, Recombinant

T4 DNA Polymerase

DNA Endonucleases

Baseline-ZERO™ DNase

Endonuclease IV, *E. coli*

T4 Endonuclease V

Lambda Terminase

RNase-Free DNase I

Pvu Rts1I Endonuclease

RNA Exonucleases

RNase R

Terminator™ 5′-Phosphate-Dependent Exonuclease

RNA Endonucleases

RNase A

RNase I, *E. coli*

RNase III, *E. coli*

RNase H, *E. coli*

Hybridase™ Thermostable RNase H

RNase T1, *Aspergillus oryzae*

RiboShredder™ RNase Blend

RNase A

DNA Ligases

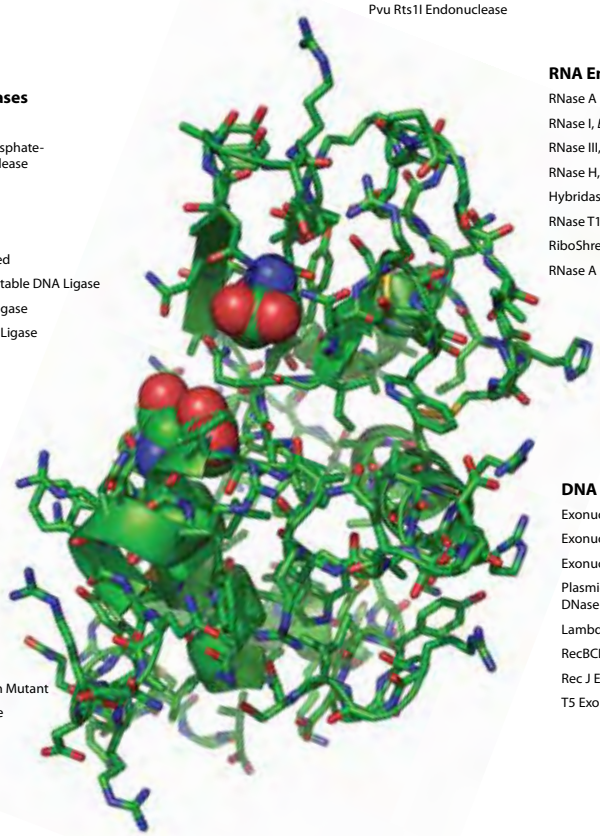
T4 DNA Ligase, Cloned

Ampligase® Thermostable DNA Ligase

CircLigase™ ssDNA Ligase

CircLigase™ II ssDNA Ligase

E. coli DNA Ligase



RNA Polymerases

T7 RNA Polymerase

T7 R&DNA™ Polymerase

DNA Exonucleases

Exonuclease I, *E. coli*

Exonuclease III, *E. coli*

Exonuclease VII

Plasmid-Safe™ ATP-Dependent DNase

Lambda Exonuclease

RecBCD Nuclease, *E. coli*

Rec J Exonuclease

T5 Exonuclease

RNA Ligases

T4 RNA Ligase

T4 RNA Ligase 2, Deletion Mutant

Thermostable RNA Ligase

Phosphatases/Kinases

APex™ Heat-Labile Alkaline Phosphatase

Tobacco Acid Pyrophosphatase (TAP)

RNA 5′ Polyphosphatase

T4 Polynucleotide Kinase, Cloned

www.epicentre.com • epicentral.blogspot.com • Toll-Free U.S. (800) 284-8474 • Fax (608) 258-3089

From the power of the HiSeq X to the speed of MiSeq, Illumina has the sequencer that's just right for you.

	 MiSeq Focused power. Speed and simplicity for targeted and small genome sequencing.	 NextSeq 500 Flexible power. Speed and simplicity for everyday genomics.		 HiSeq 2500 Production power. Power and efficiency for large-scale genomics.		 HiSeq X* Population power. $1,000 human genome and extreme throughput for population-scale sequencing.
Key applications	Small genome, amplicon, and targeted gene panel sequencing.	Everyday genome, exome, transcriptome sequencing, and more.		Production-scale genome, exome, transcriptome sequencing, and more.		Population-scale human whole-genome sequencing.
Run mode	N/A	Mid-Output	High-Output	Rapid Run	High-Output	N/A
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1 or 2
Output range	0.3-15 Gb	20-39 Gb	30-120 Gb	10-180 Gb	50-1000 Gb	1.6-1.8 Tb
Run time	5-65 hours	15-26 hours	12-30 hours	7-40 hours	< 1 day - 6 days	< 3 days
Reads per flow cell†	25 Million‡	130 Million	400 Million	300 Million	2 Billion	3 Billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp	2 × 125 bp	2 × 150 bp

* Specifications shown for an individual HiSeq X System. HiSeq X is only available as part of the HiSeq X Ten.
† Clusters passing filter.
‡ For MiSeq V3 Kits only.

Use our Sequencing Platform Comparison Tool to find the right sequencing system for your needs.

Illumina • 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. 073-2014-001 Current as of 29 May 2014

illumina®