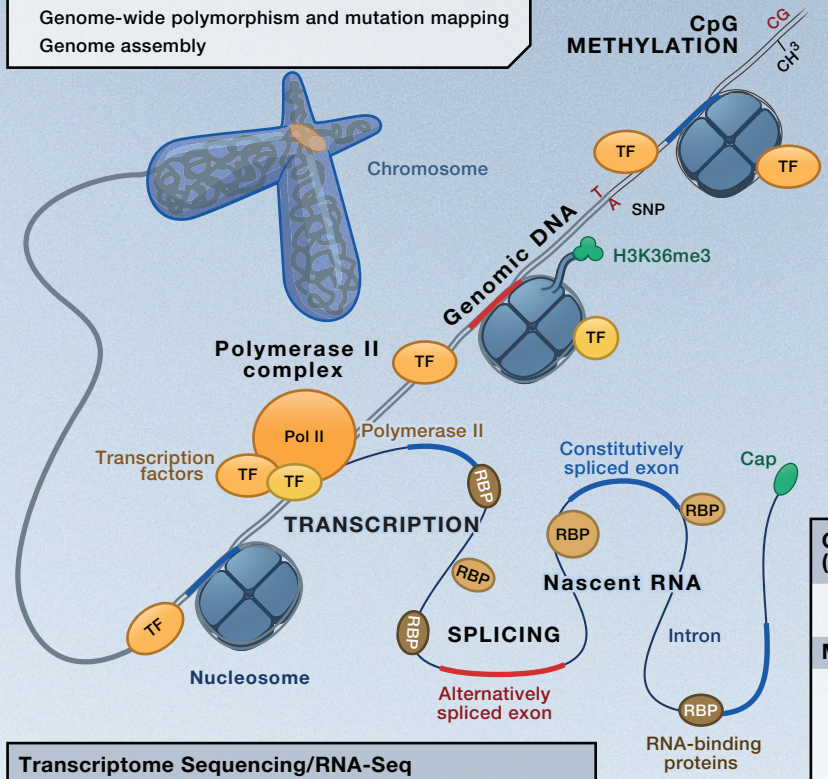
SnapShot: High–Throughput Sequencing Applications

Cell

Hong Han,[1] Razvan Nutiu,[1] Jason Moffat,[1] and Benjamin J. Blencowe[1]
[1]Banting and Best Department of Medical Research, University of Toronto, Toronto, ON M5S 3E1, Canada

Genome Sequencing/Resequencing

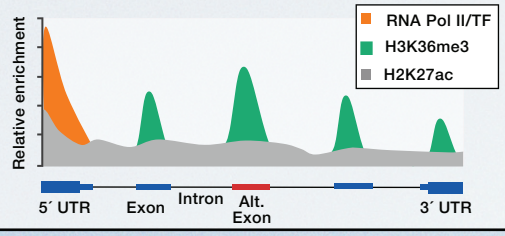
Genome-wide polymorphism and mutation mapping
Genome assembly



Chromatin Immunoprecipitation Sequencing (ChIP-Seq)

Nucleosome component, Transcription factor (TF), RNA polymerase II (Pol II) occupancy
Histone methylation or acetylation

Methyl-Seq/Bisulfite-Seq (DNA methylation status)
DNase-Seq (DNase hypersensitivity)

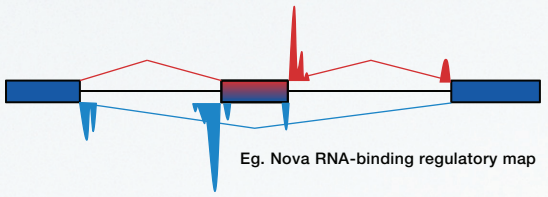


Crosslinking Immunoprecipitation Sequencing (CLIP-Seq/HITS-CLIP)

Transcriptome-wide RNA-binding protein (RBP) maps

Modified Clip For Site-specific Crosslinking

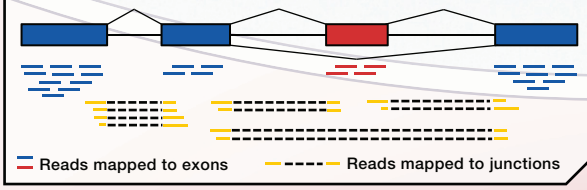
Photoactivatable ribonucleoside-enhanced CLIP (PAR-CLIP)
Individual nucleotide resolution CLIP (iCLIP)



Transcriptome Sequencing/RNA-Seq

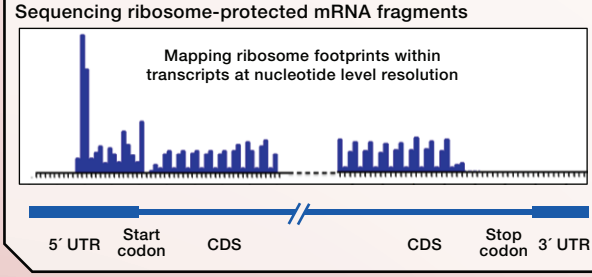
Total RNA, total RNA minus rRNA, poly(A)-selected RNA
- Gene expression profiling
- Long noncoding RNA profiling
- Alternative splicing and *trans*-splicing profiling
- Alternative polyadenylation profiling
- Mapping transcription initiation sites
- Mapping RNA editing sites (coupled with DNA-Seq)

Targeted RNA-Seq, Direct RNA-Seq, Strand-specific RNA-Seq, Nascent RNA-Seq



Ribosome Profiling

Sequencing ribosome-protected mRNA fragments



Small RNA Sequencing

e.g., microRNA/Piwi-interacting RNA profiling

Argonaute (Ago) HITS-CLIP

Mapping interactions between microRNAs and mRNAs

See online version for legend and references.

SnapShot: High–Throughput Sequencing Applications

Cell

Hong Han,[1] Razvan Nutiu,[1] Jason Moffat,[1] and Benjamin J. Blencowe[1]
[1]Banting and Best Department of Medical Research, University of Toronto, Toronto, ON M5S 3E1, Canada

Table 1. High-Throughput Next Generation Sequencing (NGS) Platforms

Platforms	Amplification; Sequencing Chemistries	Detection	Read length[a] SE-single end; PE-paired end)	Reads per lane/No. of lanes (X2 for dual flow cell)	Run time[b]
Illumina HiSeq 2000 http://www.illumina.com	Bridge Amplification; Synthesis	Fluorescence	100 bp (PE)	50×10^6 / 8 (X2)	11 days
Roche/454's GS FLX Titanium http://www.454.com	Emulsion PCR; Synthesis	Luminescence	400 bp (SE)	1×10^6	10 hr
Life/APG's SOLiD 3 www.appliedbiosystems.com	Emulsion PCR; Ligation	Fluorescence	50 bp (PE)	40×10^6 / 8 (X2)	14 days
Polonator G.007 http://www.polonator.org	Emulsion PCR; Ligation	Fluorescence	13 bp (PE)	10×10^6 / 8 (X2)	4 days
Helicos BioSciences HeliScope http://www.helicosbio.com	No amplification; Synthesis	Fluorescence	35 bp (SE)	20×10^6 / 25 (X2)	8 days
Pacific Biosciences http://www.pacificbiosciences.com	No amplification; Synthesis	Fluorescence	>1000 bp (SE)	150,000 per SMRT cell	N/A
Ion Torrent http://www.iontorrent.com	Emulsion PCR; Synthesis	Change in pH	200 bp (SE)	Variable	<2 hr
Complete Genomics Analysis Platform http://www.completegenomics.com	DNA nanoballs; Ligation	Fluorescence	Complete genomic analysis service at 40× human genome coverage; >90% of the full genome resolved (both alleles)		400 human genomes per month

Modified and updated from Metzker (2010).

[a]Average read length. [b]Run time for full sequencing experiment.
Sequencing technologies in development: Nanopore sequencing (Oxford Nanopore: http://www.nanoporetech.com; Nabsys: http://www.nabsys.com). Electron Microscopy base sequencing (Halcyon Molecular: http://halcyonmolecular.com; ZS Genetics: http://ww.zsgenetics.com).

High-throughput, next-generation sequencing (NGS) technologies have revolutionized genomics, epigenomics and transcriptomics studies by allowing massively parallel sequencing at a relatively low cost. In this SnapShot, we highlight the increasingly diverse applications of NGS, including genome sequencing/resequencing, transcriptome sequencing, small RNA sequencing, analysis of DNA/RNA-protein interactions, and ribosome profiling. In addition, we provide a quick guide (Table 1) to the currently available NGS platforms, together with their underlying methodologies and unique features.

Genome Sequencing/Resequencing

Whole-genome sequencing/resequencing and targeted genome resequencing have been used extensively for sequence polymorphism discovery and mutation mapping. These applications are rapidly advancing our understanding of human health and disease and are also facilitating the de novo assembly of uncharacterized genomes.

DNA-Protein Interactions and Epigenome Sequencing

Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) is a powerful technique for genome-wide profiling of DNA-protein interactions and epigenetic marks. It has facilitated a wide range of biological studies, including transcription factor binding, RNA polymerase occupancy, nucleosome positioning and histone modifications. Complementary methods being used to study chromatin structure and composition are Methyl-Seq and DNase-Seq for profiling DNA methylation and DNase-hypersensitive sites, respectively.

Transcriptome Sequencing/RNA-Seq

The introduction of transcriptome sequencing/RNA-Seq has provided a new approach for characterizing and quantifying transcripts. In general, total RNA, rRNA-depleted total RNA, or poly(A)-selected RNA are converted to double-stranded cDNA fragments that are then subjected to high-throughput sequencing. This strategy has been applied for profiling mRNA and noncoding RNA expression, alternative splicing, *trans*-splicing, and alternative polyadenylation and for mapping transcription initiation, termination, and RNA editing sites. Related applications include targeted RNA-Seq, direct RNA-Seq, strand-specific RNA-Seq, and nascent RNA-Seq (e.g., global run-on sequencing, GRO-Seq, and native elongating transcript sequencing, "NET-Seq").

RNA-Protein Interactions

CLIP-Seq, also known as HITS-CLIP, is a method employing in vivo crosslinking of RNA to protein followed by immunoprecipitation and high-throughput RNA sequencing to generate transcriptome-wide RNA-protein interaction maps. Modified CLIP-Seq technologies, such as PAR-CLIP (photoactivatable ribonucleoside-enhanced CLIP) and iCLIP (individual nucleotide resolution CLIP), have been applied to increase crosslinking efficiency and resolution.

Small RNA Sequencing

Similar to RNA-Seq, sequencing of size-selected short RNA provides insight into small RNA populations in different organisms, tissue and cell types, developmental stages, and disease states. It has greatly contributed to our understanding of the functions and regulatory mechanisms of different classes of small RNAs, such as microRNAs (miRNAs) and Piwi-interacting RNAs (piwiRNAs). With the recent development of Argonaute (Ago) HITS-CLIP, it is possible to simultaneously detect Ago-bound microRNAs and mRNA segments, which enables the large-scale mapping of in vivo miRNA-mRNA interactions.

Ribosome Profiling

In addition to the profound impact of NGS on transcriptomic studies, the development of methods enabling high-throughput sequencing of ribosome-protected mRNA fragments has provided a powerful tool for the analysis of translationally engaged mRNA on a genome-wide scale.

Additional Applications and Future Directions

High-throughput sequencing is a rapidly evolving technology and will likely continue to change the face of "omics" studies in the years to come. Although NGS technologies power a wide spectrum of current research applications, new innovations are continually being developed. These include barcode sequencing strategies for multiplexing the analysis of samples, metagenomic analyses, protein-protein interactome mapping ("Stitch-Seq"), and high-definition measurement of DNA-affinity landscapes (HiTS-FLIP). Future technical advances and applications are expected to further revolutionize our understanding of evolutionary biology and genotype-phenotype relationships and ultimately to bring personalized medicine into the clinic.

SnapShot: High–Throughput Sequencing Applications

Cell

Hong Han,[1] Razvan Nutiu,[1] Jason Moffat,[1] and Benjamin J. Blencowe[1]
[1]Banting and Best Department of Medical Research, University of Toronto, Toronto, ON M5S 3E1, Canada

REFERENCES

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.

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.

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.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464–469.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nat. Rev. Genet. *11*, 31–46.

Nutiu, R., Friedman, R.C., Luo, S., Khrebtukova, I., Silva, D., Li, R., Zhang, L., Schroth, G.P., and Burge, C.B. (2011). Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. Nat. Biotechnol. *29*, 659–664.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413–1415.

Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat. Rev. Genet. *10*, 669–680.

Smith, A.M., Heisler, L.E., St Onge, R.P., Farias-Hesson, E., Wallace, I.M., Bodeau, J., Harris, A.N., Perry, K.M., Giaever, G., Pourmand, N., and Nislow, C. (2010). Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. Nucleic Acids Res. *38*, e142.

Yu, H., Tardivo, L., Tam, S., Weiner, E., Gebreab, F., Fan, C., Svrzikapa, N., Hirozane-Kishikawa, T., Rietman, E., Yang, X., et al. (2011). Next-generation sequencing to generate interactome datasets. Nat. Methods *8*, 478–480.

  DOI 10.1016/j.cell.2011.09.002