

# **Sequencing**

## **Part I: Technologies & Applications**

by

**Prof. K. K. Jain**  
**MD, FRACS, FFPM**  
**Jain PharmaBiotech**  
**Basel, Switzerland**

**November 2021**

**A Jain PharmaBiotech Report**

## A U T H O R ' S   B I O G R A P H Y

Professor K. K. Jain is a neurologist/neurosurgeon with specialist qualifications including Fellowships of the Royal Colleges of Surgeons in Australia and Canada. He has trained, practiced, and held academic positions in several countries including Switzerland, India, Iran, Germany Canada, and USA. After retirement from neurosurgery, Prof. Jain remains a consultant in neurology. He is also working in the biotechnology/biopharmaceuticals industry and is a Fellow of the Faculty of Pharmaceutical Medicine of the Royal College of Physicians of UK. Currently, he is the CEO of Jain PharmaBiotech.

Prof. Jain's 492 publications include 35 books (6 as editor+ 29 as author) and 50 special reports, which have covered important areas in biotechnology, gene therapy and biopharmaceuticals, biomarkers: proteomics, molecular diagnostics, nanobiotechnology, and personalized medicine. Contributions to MedLink, an accredited continuing education program for neurologists, include 172 articles out of a total of 1250 articles by 450 authors. These articles are updated on a yearly basis. Prof. Jain's earlier books were the first in the areas covered: "Handbook of Laser Neurosurgery" (Charles C. Thomas, Springfield, Ill, 1983) and "Textbook of Hyperbaric Medicine" (1st ed in 1990 and 6th ed by Springer, 2017). His "Textbook of Gene Therapy" was translated into Chinese in 2000. Recent books include "Handbook of Nanomedicine" (Springer 2008, Chinese edition by Peking University Press 2011, 3rd ed 2017), "Textbook of Personalized Medicine" (Springer 2009; Japanese ed 2012; 2nd ed Springer 2015, 3rd ed 2021), "Handbook of Biomarkers" (Springer 2010; Chinese ed, Chemical Industry Press 2016, 2nd ed 2017), "Drug-induced Neurological Disorders", 4th ed (Springer 2021), "Handbook of Neuroprotection" (Springer 2011, 2nd ed 2019), "Applications of Biotechnology in Cardiovascular Therapeutics" (Springer 2011), "Applications of Biotechnology in Neurology" (Springer 2013), and "Applications of Biotechnology in Oncology" (Springer 2014). He has also edited 3 editions of "Drug Delivery System" (Springer 2008, 20012 and 2020) and "Applied Neurogenomics" (Springer 2015). Lectures on personalized medicine given at Kazakh National Medical University, Kazakhstan were translated into Russian and published as a book "Essentials of Personalized Medicine" (LITERRA Publishing House, Moscow, 2019). Currently, he is writing "The Handbook of Alzheimer Disease" to be published by Springer in 2022.

## A B O U T   T H I S   R E P O R T

This report will cover sequencing technologies, equipment, reagents, markets, and companies involved in this area. It is an extension of the report on Molecular Diagnostics, which has been in continuous publication since 1995. New advances and applications of sequencing have justified the issuance of a separate report on the topic.

**November 2021**  
**Copyright © 2021 by**

**Jain PharmaBiotech**  
Bläsiring 7  
CH-4057 Basel  
Switzerland

**Tel & Fax:** +4161-6924461  
**Email:** info@pharmabiotech.ch  
**Web site:** http://pharmabiotech.ch/

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise without the prior written permission of the Publisher. This report may not be lent, resold or otherwise traded in any manner without the consent of the Publisher. While all reasonable steps have been taken to ensure the accuracy of the information presented, the Publisher cannot accept responsibility for inadvertent errors or omissions.

# T A B L E   O F   C O N T E N T S

<b>0. Executive Summary .....</b>	<b>16</b>
<b>1. Introduction .....</b>	<b>18</b>
<b>Definition and scope of sequencing .....</b>	<b>18</b>
<b>Historical aspects of sequencing.....</b>	<b>18</b>
<b>Basics of molecular biology .....</b>	<b>19</b>
DNA .....	19
<i>DNA polymerases.....</i>	20
<i>DNA helicases.....</i>	20
<i>Restriction endonucleases.....</i>	20
RNA .....	20
<i>RNA polymerases.....</i>	21
<i>Non-coding RNAs.....</i>	21
DNA transcription.....	21
Chromosomes .....	22
<i>Chromatin.....</i>	22
<i>Chromosome sequences .....</i>	22
Telomeres.....	23
Mitochondrial DNA .....	23
Genes.....	24
<i>The genetic code.....</i>	24
<i>Gene expression.....</i>	24
<i>Genes and human diseases.....</i>	25
The human genome .....	25
<i>ENCODE .....</i>	26
The epigenome.....	27
<i>Epigenetics and epigenomics .....</i>	27
<i>DNA methylation.....</i>	27
<i>Human Epigenome Atlas .....</i>	28
Metagenomics .....	28
<b>Variations in the human genome .....</b>	<b>28</b>
Variations in DNA sequences .....	29
Single nucleotide polymorphisms .....	29
Haplotyping.....	30
Complex chromosomal rearrangements .....	30
Insertions and deletions in the human genome.....	30
Large scale variation in human genome .....	31
Variation in copy number in the human genome .....	31
Structural variations in the human genome .....	33
Transposons.....	34
<i>Retrotransposon capture sequencing .....</i>	34
Mapping and sequencing of structural variation from human genomes.....	35
<b>Impact of sequencing on healthcare .....</b>	<b>35</b>
<b>2. DNA Sequencing Technologies.....</b>	<b>38</b>
<b>Introduction .....</b>	<b>38</b>
DNA extraction and sample preparation .....	39
<i>Apollo 300 System for next generation sequencing .....</i>	40
<i>Electrophoresis-based method .....</i>	40
<i>Ion OneTouch System .....</i>	41
<i>Microfluidics-based extraction and sample preparation.....</i>	41
<i>Pressure Cycling Technology .....</i>	41
<i>Selective immobilization of nucleic acids onto magnetic microparticles.....</i>	41
<i>Targeted and hybridization-based DNA capture .....</i>	41
Sanger-sequencing technology .....	42
<i>Dye-terminator sequencing .....</i>	42
<i>Large-scale sequencing.....</i>	43
<i>Automated DNA-sequencing .....</i>	43
<i>Enhancements of Sanger-sequencing .....</i>	43
<i>ABI PRISM® 310 Genetic Analyzer.....</i>	44
<i>Life Technologies' 3500 Dx genetic analyzer .....</i>	44
<b>Limitations of sequencing methods and measures to remedy them.....</b>	<b>44</b>
<i>Paired end transcriptome sequencing to overcome short read lengths .....</i>	45
<i>Long vs short read lengths .....</i>	45
<i>Validation of NGS data.....</i>	46
<b>Emerging sequencing technologies.....</b>	<b>46</b>
Chemical DNA sequencing .....	47

Chemical affinity capture and massively parallel DNA sequencing.....	48
<b>Next generation sequencers .....</b>	<b>49</b>
4300 DNA analyzer .....	49
Apollo 100 .....	49
Life Technologies 3500 series Genetic Analyzer.....	49
Cyclic array sequencing .....	50
CEQ™ 8000.....	50
DeepCAGE sequencing.....	50
Electron microscope-based DNA sequencing.....	51
GS-FLEX system (Roche) .....	51
<i>Background of sequencing technology</i> .....	51
IBS sequencing technology .....	54
Illumina Genome Analyzer System .....	55
<i>Next generation sequencing based on Lightning Terminators</i> .....	56
<i>MiSeq FGx</i> .....	57
<i>NovaSeq 6000 system</i> .....	57
Ion Torrent's sequencing technology .....	57
<i>Ion S5 system</i> .....	58
MegaBACE 500 .....	58
Microdroplet-based PCR for large-scale targeted sequencing.....	59
Millikan sequencing .....	59
Multiplex amplification of human DNA sequences .....	59
Nanoscale sequencing .....	60
Polonator sequencer.....	60
RainStorm™ microdroplet technology .....	61
Sequential DEXAS.....	61
SOLiD system: sequencing by ligation .....	62
PCR-based DNA sequencing technologies .....	62
<i>Bridge amplification PCR system</i> .....	63
<i>COLD-PCR and sequencing</i> .....	63
<i>Digital PCR</i> .....	64
<i>Dual primer emulsion PCR</i> .....	64
<i>Emulsion PCR</i> .....	65
<i>Multiplex PCR</i> .....	65
<i>SiMSen-seq for ultrasensitive detection of mutations</i> .....	65
Non-PCR based sequencing .....	66
Nucleic acid sequence-based amplification .....	66
Microarray-based DNA sequencing technologies .....	66
<i>ArrayIt's® H25K</i> .....	67
<i>High-throughput array-based resequencing</i> .....	67
<i>Sequencing by hybridization</i> .....	67
<i>SOLiD-System based ChIP-Sequencing</i> .....	68
<i>Companies developing whole genome chips/microarrays</i> .....	69
Next generation sequencing vs microarrays for gene expression profiling.....	69
RNA sequencing .....	70
<i>RNA-seq vs microarrays</i> .....	71
<i>Capture sequencing</i> .....	71
<i>ChIP-Seq</i> .....	71
<i>In situ sequencing</i> .....	72
<i>Strand specific RNA sequencing</i> .....	72
<i>SHAPE-Seq</i> .....	73
<i>Applications of RNA-seq</i> .....	73
<i>Challenges for future development of RNA-seq</i> .....	74
Exome sequencing .....	74
<i>Limitations of exome sequencing</i> .....	74
<i>Human exome microarrays</i> .....	75
<i>WES vs WGS</i> .....	75
Third generation sequencing .....	76
<i>SOLiD4 System</i> .....	76
<i>SOLiD PI System</i> .....	77
<i>Helicos™ Genetic Analysis System</i> .....	77
<i>Molecular Combing</i> .....	78
Nanotechnology-based sequencing .....	78
<i>DNA sequence by use of nanoparticles</i> .....	78
<i>Denaturation mapping of DNA in nanofluidic channels</i> .....	79
<i>DNBseq technology</i> .....	79
<i>Nanopore sequencing</i> .....	79
<i>Sequencing through graphene nanopores</i> .....	82
Convex lens-induced nanoscale templating .....	82
Detection of single molecules for sequencing .....	83
<i>Helicos™ Genetic Analysis System</i> .....	84

<i>Molecular Combing</i>	85
<i>Optical Mapping</i>	85
<i>Nanopore-based single-molecule detection of specific DNA sequences</i>	86
<i>Phasing through the sequencing of single molecules</i>	87
<i>Sequencing-by-synthesis for single-molecule sequencing</i>	87
<i>Single molecule DNA sequencing by use of carbon nanotubes</i>	88
<i>Single molecule sequencing using Qdot nanocrystals</i>	88
<i>Single-molecule DNA sequencing in a sTOP chip nanowell</i>	89
<i>Single-molecule real-time sequencing</i>	90
<i>Single molecule targeted sequencing using GenoCare™ Analyzer</i>	91
<b>Single cell sequencing</b>	91
<i>Drop-seq for single cell sequencing</i>	91
<i>MALBAC for single cell DNA sequencing</i>	92
<i>Microfluidic single-cell whole-transcriptome sequencing</i>	92
<i>Optical-mapping and whole-genome sequencing of cells</i>	93
<i>Single cell sequencing of uncultured microbes</i>	93
<i>Single sperm sequencing</i>	93
<i>Single cell RNA sequencing</i>	94
<i>Single-cell combinatorial indexed sequencing</i>	95
<i>Single-stranded sequencing using microfluidic reactors</i>	95
<i>Future of single cell sequencing</i>	95
<i>Thermosequencing</i>	96
<i>Whole genome sequencing for haplotyping</i>	96
<i>Mitochondrial exome sequencing</i>	97
<i>ImmunoSEQ technology</i>	98
<b>Future prospects of next generation sequencing</b>	98
Devices for NGS	98
Reduction of errors in NGS	99
<i>Artifactual mutations during the sample preparation process</i>	99
<i>Contamination in high throughput sequencing</i>	99
<i>Duplex sequencing</i>	99
<b>3. Role of Bioinformatics in Sequencing</b>	102
<b>Introduction</b>	102
<b>Growth of the sequencing database</b>	102
<b>Sequencing data storage</b>	102
Cloud computing for sequencing data	103
<i>Cloud computing facilities</i>	103
<b>Bioinformatics challenges of new sequencing technology</b>	103
<b>Construction of libraries for NGS</b>	104
<b>Bioinformatic tools for analysis of genomic sequencing data</b>	105
Software for DNA sequencing	106
<i>Software from academic and open sources</i>	106
<i>Commercial software for sequencing</i>	107
ChIA-PET tool for analysis with paired-end tag sequencing	109
CLC Cancer Research Workbench	109
Compressive genomics	109
Detection of CNVs and gene duplications	110
Detection of SVs in massively parallel sequencing data	110
Differential expression analysis for sequence count data	111
DRAGEN™ Bio-IT Processor	111
Expression profiling without genome sequence information	111
Flye long-read assembly algorithm	112
Ingenuity® Variant Analysis™	112
Ion Reporter Software	112
Opal platform	113
VAAST	113
<b>Accessing DNA sequence information</b>	113
Clinical Genomicist Workspace for managing NGS-based clinical tests	113
<b>Analysis of genomic variation by sequencing of large populations</b>	114
Analysis of rare variants in NGS studies	114
<b>Human gene connectome</b>	115
<b>Funding of research for interpretation of sequencing data</b>	115
<b>Future challenges for managing sequencing data</b>	116
<b>4. Comparative Analysis of Sequencing Technologies</b>	118
<b>General findings of the study</b>	118
Sanger versus second generation marketed sequencers	118
Common features and differences among second generation sequencers	118
<b>Third generation large sequencers</b>	119
SOLiD4 versus competing large sequencers	119

Illumina's HiSeq sequencer .....	120
<b>Third generation desktop sequencers .....</b>	<b>121</b>
BGI's BGISEQ-500 desk top sequencer .....	121
Illumina's MiSeq sequencer .....	121
Roche GS Junior System .....	122
Life Technologies' Benchtop Ion Proton™ Sequencer .....	122
Oxford Nanopore's MinION™ system .....	123
<b>The ideal desk-top sequencer .....</b>	<b>123</b>
<b>SWOT analysis of small sequencers .....</b>	<b>123</b>
<b>Concluding remarks on SWOT analysis of desktop sequencers .....</b>	<b>127</b>
 <b>5. Sequencing for Research .....</b>	<b>130</b>
<b>Introduction .....</b>	<b>130</b>
<b>Applications in basic research.....</b>	<b>131</b>
Barcode sequencing of synapses in the brain .....	131
ChIA-PET technology for 3D study of the genome .....	131
ChIP-Seq for study of gene expression .....	132
Chromatin profiling by direct DNA sequencing .....	132
Discovery of immunoglobulin gene by pyrosequencing.....	132
Epigenetic modifications analyzed by next generation sequencing .....	133
Exome sequencing for study of human variation .....	133
Genome sequencing with combinatorial probe anchor ligation.....	134
Identifying protein-coding genes in genomic sequences .....	134
Mutation rate measured by direct sequencing .....	135
Protein-protein interactome network mapping .....	136
RNA sequencing.....	136
<i>Applications of RNA sequencing .....</i>	136
<i>Contribution of RNA structure to gene regulation revealed by RNA-Seq.....</i>	136
<i>Molecular indexing for quantitative targeted RNA sequencing .....</i>	137
Sequencing for the study of microchimerism .....	137
Sequencing for the study of CNVs .....	138
<i>CNVnator .....</i>	138
<i>GS-FLX sequencing for simultaneous detection of mutations and CNVs .....</i>	139
Sequencing the transcriptomes of stem cells .....	139
<i>Sequencing, stem cells and neurodegeneration .....</i>	139
<i>Sequencing, stem cells and regeneration .....</i>	139
Sequencing and synthetic biology .....	140
<i>Synthetic sequence in a bacterial cell .....</i>	140
<i>Synthesizing long DNA molecules.....</i>	140
<i>Functional synthetic proteins .....</i>	141
Single-cell sequencing for study of cell lineage .....	142
<b>Sequencing of human genomes .....</b>	<b>142</b>
100,000 Genomes Project .....	142
Genome of the Netherlands.....	143
Missing human genome sequences.....	144
Personal genome sequencing.....	144
Role of sequencing in identification of human remains .....	145
Saudi Arabian human genome program .....	145
Sequence map of the human pan-genome .....	145
Sequencing of African genomes .....	146
Sequencing of Korean genomes .....	146
Sequencing mitochondrial genome .....	147
Sequencing of ancient genomes.....	147
<i>Ancient DNA evidence for genomic variations in South East Asia .....</i>	147
<i>Anzick-1 genome .....</i>	148
<i>Genome of the Kennewick Man .....</i>	148
<i>Hominin genome.....</i>	148
<i>Indus valley civilization and DNA sequencing .....</i>	149
<i>Neandertal genome.....</i>	149
<i>Saqqara genome.....</i>	151
<i>Sequencing and controversies about origin of Native Americans.....</i>	151
Whole genome sequencing .....	152
Whole-genome sequencing of methylome .....	152
Whole genome resequencing .....	153
Prospects of genome sequencing for human ancestry .....	153
<b>Sequencing genomes of non-human primates .....</b>	<b>153</b>
Sequencing of chimpanzee genome.....	153
Sequencing of macaque genome.....	154
Sequencing of gorilla genome.....	154
<b>Sequence the genomes of vertebrate species .....</b>	<b>155</b>
<b>Sequencing genomes of other organisms .....</b>	<b>155</b>

Species biodiversity study by the Canadian Centre for DNA Barcoding .....	156
Ant genome .....	156
Avian genomes.....	157
Bat genome .....	157
Body louse genome.....	158
Camel genome .....	159
Cat genome .....	159
Dog genome.....	160
Frog genome.....	161
Goat genome .....	161
Horse genome .....	162
<i>Ancient horse genome</i> .....	163
House fly .....	163
Mosquito genome.....	163
Mouse genome .....	164
Tibetan antelope genome.....	164
Turkey genome .....	164
Water flea genome.....	165
Whale genome .....	166
Woolly mammoth genome .....	166
<b>Sequencing studies of the human microbiome.....</b>	<b>167</b>
Human Microbiome Project.....	167
Human virome.....	168
Microbial codes to distinguish individuals .....	168
<b>Single cell sequencing of the brain .....</b>	<b>168</b>
<b>Population targeted sequencing studies .....</b>	<b>169</b>
<b>Transcriptome sequencing for mRNA expression.....</b>	<b>170</b>
<b>RNA splice variants.....</b>	<b>170</b>
<b>Sequencing projects supported by US Government.....</b>	<b>171</b>
NHGRI's sequencing initiatives.....	171
JGI's Community Sequencing Program .....	171
NIH funding for interpreting sequence variants in the human genome .....	172
NIH to fund studies of gene-environmental interactions in human diseases .....	172
Approved medical sequencing projects .....	173
<b>1000 Genomes Project.....</b>	<b>174</b>
Findings of some studies of the 1000 genomes project .....	174
HapMap catalog as a foundation .....	175
Role of SOLiD™ System in 1000 Genomes Project.....	175
Protection of privacy of participants in 100 Genomes project .....	175
Concluding remarks .....	176
<b>Human Variome Project .....</b>	<b>176</b>
<b>Sequencing in space .....</b>	<b>177</b>
<b>Academic centers conducting research on sequencing .....</b>	<b>178</b>
Important academic collaborations .....	179
Manpower for sequencing .....	180
New York Genome Center .....	180
USTAR Center for Genetic Discovery.....	181
US Government funding for research .....	181
<i>Joint Genome Institute</i> .....	181
<i>NHGRI's Clinical Sequencing Exploratory Research grants</i> .....	182
<b>6. Applications of Sequencing in Healthcare .....</b>	<b>184</b>
<b>Introduction .....</b>	<b>184</b>
<b>Applications of sequencing in molecular diagnostics .....</b>	<b>184</b>
Clinical exome sequencing .....	185
<i>ACE Clinical Exome test</i> .....	186
Diagnosis and screening of genetic disorders.....	186
<i>Cystic fibrosis screening using high-throughput NGS</i> .....	186
<i>CNV sequencing for diagnosis of chromosomal disorders</i> .....	187
<i>Genetic misdiagnosis due to limited population sequencing data</i> .....	187
<i>Karyotyping based on NGS</i> .....	187
<i>NGS for diagnosis of rare genetic disorders</i> .....	188
<i>Role of WGS in screening of newborns</i> .....	188
<i>Role of WGS in chromothripsis</i> .....	188
<i>Single cell sequencing for PGD</i> .....	189
<i>WGS for identification of genetic disorders in critically ill infants</i> .....	189
<i>WGS for pre-implantation genetic diagnosis in IVF</i> .....	189
Guidelines for use of sequencing for diagnosis .....	190
Incidental findings on clinical sequencing .....	191
NGS for diagnosis of CNS infections .....	191
NGS for detection of solid organ transplant rejection .....	191

NGS for forensic diagnosis .....	192
NGS-based molecular diagnosis at POC using mobile phone microscopy.....	192
Companies developing sequence-based molecular diagnostics .....	193
<b>Sequencing in cardiovascular disorders.....</b>	<b>194</b>
Inherited cardiomyopathies .....	194
Exome sequencing and mutations associated with risk of coronary heart disease .....	195
<b>Sequencing for study of the human immune system.....</b>	<b>195</b>
Sequencing for investigating drug-virome interactions in organ transplants.....	196
Immune profiling using a synthetic human virome .....	197
NGS-based HLA typing .....	197
<i>High-throughput HLA genotyping with deep sequencing.....</i>	198
<i>Approaches to NGS-based HLA typing.....</i>	198
<i>Applications of NGS-based HLA typing .....</i>	199
<b>Applications of sequencing relevant to the human microbiome .....</b>	<b>199</b>
Sequencing the gut microbiome for discovery of immunomodulators .....	200
Sequencing of gut microbes in obesity.....	200
Companies developing microbiome-based products .....	200
<b>Pharmaceutical applications of sequencing .....</b>	<b>201</b>
Drug discovery and development .....	201
Resequencing .....	202
RNA profiling .....	202
Quantitative selection of aptamers through sequencing .....	203
Next generation sequencing and drug safety .....	203
<b>Sequencing in aging research .....</b>	<b>204</b>
WGS for the study of supercentenarians .....	204
<i>Sequencing of APOE gene in centenarians.....</i>	204
<b>7. Applications of Sequencing in Oncology .....</b>	<b>206</b>
<b>Introduction .....</b>	<b>206</b>
<b>Sequencing technologies for cancer.....</b>	<b>206</b>
A project to assess sequencing technologies for tumor DNA .....	206
A universal NGS-based oncology test system .....	207
Amplicon sequencing in cancer .....	207
ASCO guidelines for use of NGS to test cancer susceptibility.....	207
Cancer Genome Atlas .....	208
Catalog of cancer genes.....	208
Detection of cancer biomarkers.....	209
<i>Sequencing mitochondrial DNA to identify cancer biomarkers .....</i>	209
<i>Biomarkers for personalizing cancer treatment .....</i>	210
Digital proteomics for cancer profiling.....	211
Discrepancies in cancer genome sequencing.....	211
Epigenome profiling .....	211
Exosome sequencing .....	212
Gaining insights into mutational processes .....	212
Multiple-gene sequencing panel to assess risk of hereditary cancer .....	212
Multiplexed cancer gene mutation analysis.....	212
NGS for enhancing verification rate of chromosomal structural rearrangements .....	213
NGS-based molecular profiling of cancer in FFPE specimens .....	214
Paired-end sequencing .....	215
Pathology tissue-ChIP.....	216
Quality control of NGS in oncology .....	216
RNA-Seq to study cancer transcriptome .....	216
Sequencing cancer cell lines .....	217
<b>Sequencing for studying somatic mutations in cancer .....</b>	<b>217</b>
Chromothripsis in cancer .....	217
Microsatellite instability in cancer genome.....	218
Somatically acquired genomic rearrangements in cancer.....	218
Sequencing for identification of FGFR gene fusions in cancer .....	218
Sequencing single cells to study evolution of cancer .....	219
Sequencing circulating tumor cell genomes .....	219
<b>Sequencing for detection of germline cancer risk variants .....</b>	<b>220</b>
<b>High throughput sequencing for anticancer drug discovery .....</b>	<b>221</b>
<b>NGS for developing targeted cancer therapies .....</b>	<b>221</b>
<b>Sequencing for assessing resistance to anticancer therapy .....</b>	<b>221</b>
<b>Sequencing of tumors of various organs.....</b>	<b>222</b>
Brain tumors .....	222
<i>Sequencing for genetic alterations in gliomas .....</i>	222
<i>Sequencing for genetic alterations in medulloblastoma .....</i>	222
Breast cancer .....	223
<i>BRCA mutations.....</i>	223
<i>Circulating nucleic acids as biomarkers of cancer .....</i>	224

<i>Deep sequencing of miRNA for signatures of invasiveness</i> .....	224
<i>NGS reveals heterogeneity of breast cancer</i> .....	224
<i>Sequencing of breast cancer metastases</i> .....	225
<i>Triple negative breast cancer</i> .....	225
<i>Whole genome sequencing in breast cancer</i> .....	226
Colorectal cancer .....	226
Gastric cancer .....	227
Head and neck cancer .....	228
<i>NGS for detection of HPV sequences in carcinoma of oropharynx</i> .....	228
Hematological malignancies .....	228
<i>Acute myeloid leukemia</i> .....	228
<i>Acute promyelocytic leukemia</i> .....	230
<i>Chronic myelomonocytic leukemia</i> .....	230
<i>Hairy-cell leukemia</i> .....	230
<i>Hematological cancer risk inferred from blood DNA sequence</i> .....	230
<i>Myelodysplastic syndromes</i> .....	231
<i>Sequencing in chronic neutrophilic leukemia and atypical CML</i> .....	231
Sequencing in hepatocellular carcinoma .....	232
Lung cancer .....	232
<i>NGS to distinguish primary lung cancer from pulmonary metastases</i> .....	232
<i>Sequencing of small cell lung cancer</i> .....	233
Melanoma .....	233
Ovarian cancer .....	233
Prostate cancer .....	234
<i>Identification of mutations in prostate cancer by exome sequencing</i> .....	235
<i>Role of sequencing in liquid biopsy for prostate cancer patients</i> .....	236
Current status and future of NGS applications in oncology .....	236
<i>Actionable Genome Consortium to guide NGS in clinical oncology</i> .....	237
<i>Guidelines for interpretation/reporting of sequence variants in cancer</i> .....	238
<i>Points to consider for germline findings in tumor-only sequencing</i> .....	238
<i>Translation of gene panel sequencing into routine cancer diagnosis</i> .....	239

## **8. Sequencing in Genetic Disorders ..... 241**

<b>Introduction .....</b>	<b>241</b>
<b>Approaches to sequencing in genetic disorders .....</b>	<b>241</b>
DNA sequencing for prenatal disorders .....	241
High-throughput sequencing in Undiagnosed Disease Program at NIH .....	242
Sequencing of maternal plasma for detection of fetal aneuploidy .....	242
Sequencing for study of transposons .....	244
Sequencing genomes of the newborn to screen for genetic disorders .....	244
Study of rare variants in pinpointing disease-causing genes .....	245
Tandem repeat variability for detection of genetic factors in diseases .....	246
Whole genome sequencing for diagnosis of genetic disorders .....	247
Whole exome sequencing for diagnosis of genetic disorders .....	247
WES and WGS in monogenic disorders .....	248
Whole genome sequencing of a human fetus from maternal plasma .....	249
<b>Genetic disorders investigated by sequencing .....</b>	<b>250</b>
Bartter syndrome .....	250
CHARGE syndrome .....	250
DiGeorge syndrome .....	251
Discovery of the gene for Miller syndrome .....	251
Discovery of the gene for Kabuki syndrome .....	252
Familial combined hypolipidemia .....	252
Familial thoracic aortic aneurysm .....	252
Hereditary ataxia .....	252
Hereditary blindness .....	253
Neurofibromatosis type 1 .....	254
Noonan syndrome .....	254
Proteus syndrome .....	255
Syndrome of hypogonadotropic hypogonadism, ataxia, and dementia .....	255
Syndrome of polyarthritis nodosa vasculopathy .....	255
X-linked disorder due to N-terminal acetyltransferase deficiency .....	256
Sequencing in mitochondrial disorders .....	256
<i>Chronic progressive external ophthalmoplegia</i> .....	257
Role of NGS in prevention of chromosome abnormalities & monogenic disorders .....	257

## **9. Sequencing in Neurological and Psychiatric Disorders..... 258**

<b>Introduction .....</b>	<b>258</b>
<b>Sequencing in Alzheimer disease .....</b>	<b>258</b>
<b>Sequencing in Parkinson disease .....</b>	<b>260</b>
<b>Sequencing in Huntington's disease .....</b>	<b>260</b>

<b>Sequencing in Wilson's disease .....</b>	<b>261</b>
<b>Sequencing in ataxias .....</b>	<b>261</b>
<b>Sequencing in epilepsy .....</b>	<b>262</b>
Epileptic encephalopathy.....	262
<b>Sequencing for mutations in familial amyotrophic lateral sclerosis .....</b>	<b>263</b>
<b>Sequencing of whole genome in Charcot-Marie-Tooth disease .....</b>	<b>263</b>
<b>Sequence-based detection of a variant of Lambert-Eaton syndrome .....</b>	<b>263</b>
<b>Sequencing in muscular dystrophy .....</b>	<b>264</b>
<b>Sequencing in acute brain injury due to hemorrhage.....</b>	<b>264</b>
<b>Sequencing for mutations in autism spectrum disorders .....</b>	<b>265</b>
<b>Sequencing for diagnosis of intellectual disability .....</b>	<b>265</b>
<b>Sequencing in neurodevelopmental disorders .....</b>	<b>266</b>
NGS for identifying mutations in RNA gene.....	266
WES in neurodevelopmental disorders .....	266
<b>Sequencing in Möbius syndrome.....</b>	<b>267</b>
<b>Sequencing in attention-deficit/hyperactivity disorder .....</b>	<b>267</b>
<b>Sequencing in schizophrenia and bipolar disorder .....</b>	<b>267</b>
<b>Sequencing in drug addiction.....</b>	<b>267</b>
 <b>10. Applications of sequencings in infections .....</b>	<b>270</b>
<b>Introduction .....</b>	<b>270</b>
<b>DNA sequencing for study of bacterial epidemics .....</b>	<b>270</b>
Sequencing of GAS genotype emm89 .....	271
Genome sequencing of <i>H. influenzae</i> to identify population structure.....	271
Role of sequencing in cholera epidemics .....	271
Role of sequencing in epidemic of Shiga toxin-producing <i>E. coli</i> .....	272
Sequencing study of emergence of <i>M. tuberculosis</i> in East Africa .....	273
Sequencing study of Salmonella emergence in Sub-Saharan Africa .....	273
Sequencing of ancient specimens from past epidemics .....	274
<b>Sequencing for tracking hospital acquired infections.....</b>	<b>274</b>
Sequencing for investigation of MRSA outbreaks.....	274
Role of sequencing in tracking a hospital infection of <i>K. pneumoniae</i> .....	275
Role of whole genome sequencing in identification of <i>C. difficile</i> .....	276
Role of NGS in diagnosis of infectious agents causing meningitis & encephalitis .....	276
<b>Role of sequencing in the management of bacterial infections .....</b>	<b>276</b>
Pyrosequencing of microbial flora in leg ulcers .....	276
Sequencing for mapping genomic variation in <i>Mycobacterium ulcerans</i> .....	277
<b>Sequencing in the management of antimicrobial drug resistance .....</b>	<b>277</b>
Maximum-depth sequencing.....	277
Sequencing for study of antibiotic resistance in bacteria .....	278
Sequencing for predicting the virulence of MRSA.....	278
Sequencing for detection of drug resistance in <i>Plasmodium falciparum</i> .....	278
WGS for diagnosis of drug-resistant <i>M. tuberculosis</i> .....	279
<b>Sequencing for investigation of food-borne infections.....</b>	<b>279</b>
<b>Sequencing for mapping genetic interactions in bacteria.....</b>	<b>279</b>
<b>Metagenomic sequencing of bacteria .....</b>	<b>280</b>
Sequencing of DNA from single cells of bacteria .....	280
<b>Sequencing of the fungal genomes .....</b>	<b>281</b>
<b>Sequencing of human salivary microbiome.....</b>	<b>281</b>
<b>Next generation sequencing for antibacterial therapeutic discovery .....</b>	<b>282</b>
<b>Sequencing for diagnosis of viral diseases.....</b>	<b>282</b>
High throughput sequencing for diagnosis of viral diseases .....	282
Sensitive metagenomic sequencing for detection of pathogenic human viruses .....	283
<b>Sequencing in the management of HIV/AIDS .....</b>	<b>283</b>
Long read sequencing for personalizing HIV therapy.....	283
NGS for studying neuroAIDS .....	284
Sequencing plus immunological analyses to study HIV evolution .....	284
Surveillance of drug resistance in HIV-infected individuals .....	284
<b>Sequencing in the management of Ebola virus infection .....</b>	<b>285</b>
<b>Sequencing in the management of HCV .....</b>	<b>286</b>
Sequencing for detection of a novel pegivirus associated with HCV.....	286
<b>Sequencing genomes of hemorrhagic fever viruses .....</b>	<b>287</b>
Sequencing genome of Lassa fever virus.....	287
Sequencing genome of a rhabdovirus associated with acute hemorrhagic fever .....	287
Surveillance of H1N1 influenza A virus using resequencing arrays .....	287
<b>Sequencing genomes of coronaviruses .....</b>	<b>288</b>
<b>Regulatory aspects of sequencing for diagnosis of infections .....</b>	<b>288</b>
<b>Future of pathogen sequencing .....</b>	<b>289</b>
 <b>11. Role of Sequencing in Personalized Medicine.....</b>	<b>292</b>
<b>Introduction .....</b>	<b>292</b>

<b>Technologies relevant to sequencing and personalized medicine .....</b>	<b>292</b>
Whole genome sequencing and personalized medicine.....	292
Whole exome sequencing and personalized disease risk.....	293
Large-scale deep sequencing of human genomes .....	294
<b>Personal Genome Project.....</b>	<b>294</b>
<b>Role of sequencing in personalized cancer management .....</b>	<b>295</b>
Circulating cell-free DNA sequencing for personalized cancer therapy .....	295
Companies and cancer centers using sequencing for personalized oncology .....	295
<b>Standardization of sequencing for personalized medicine .....</b>	<b>297</b>
Regulating genomic testing in the era of personalized medicine .....	297
<b>Future of sequencing and personalized medicine.....</b>	<b>297</b>
<b>12. Current Status and Future Prospects .....</b>	<b>300</b>
ACMG clinical laboratory standards for NGS.....	300
Regulatory and quality control issues of sequencers in the USA .....	300
International NGS standardization .....	301
Applications of NGS in clinical trials.....	302
Human transcriptome array in clinical trials.....	302
Challenges for clinical applications of NGS .....	302
Direct-to-consumer WGS services .....	303
Ethical aspects of sequencing .....	304
Future trends in clinical sequencing .....	305
Rare Diseases Genomes Project.....	305

## Tables

Table 1-1: Historical landmarks in DNA sequencing .....	19
Table 1-2: Genetic variations in the human genome.....	29
Table 2-1: ChIP detection platforms for sequencing .....	69
Table 2-2: Companies developing whole genome chips/microarrays.....	69
Table 2-3: Systems for single molecule sequencing.....	83
Table 3-1: Software programs for sequencing from open sources .....	106
Table 3-2: Companies providing DNA sequencing software .....	107
Table 4-1: Comparison of a generation I and generation II sequencers.....	118
Table 4-2: Similarities and differences between second generation sequencers.....	119
Table 4-3: SWOT of ABI 310 .....	124
Table 4-4: SWOT of IBS sequencing.....	124
Table 4-5: SWOT of NABsys' Hybridization-Assisted Nanopore Sequencing.....	124
Table 4-6: SWOT of 4300 DNA Analysis System Li-Cor .....	124
Table 4-7: SWOT of BGI's BGISEQ-500 desktop sequencer .....	124
Table 4-8: SWOT of Polonator .....	125
Table 4-9: SWOT of Roche GS FLEX Junior.....	125
Table 4-10: SWOT of Oxford Nanopore's MinION™ system.....	125
Table 4-11: SWOT of Ion Torrent™ Personal Genome Machine .....	125
Table 4-12: SWOT of Pacific BioSciences' single-molecule real-time sequencing.....	126
Table 4-13: SWOT of Illumina's MiSeqDx .....	126
Table 4-14: SWOT of Illumina's NextSeq 500 .....	126
Table 4-15: SWOT of QIAGEN's GeneReader™ sequencer .....	126
Table 4-16: SWOT of Bio-Rad (formerly GnuBio's) droplet-based sequencing system .....	127
Table 4-17: SWOT of 10X Genomics' GemCode .....	127
Table 5-1: Number of genes in organisms with fully sequenced genomes.....	130
Table 5-2: Approved medical sequencing projects.....	173
Table 5-3: Academic centers conducting research on DNA sequencing .....	178
Table 5-4: Distribution of scientific manpower for sequencing .....	180
Table 6-1: Companies involved in application of sequencing in molecular diagnostics .....	193
Table 6-2: Companies developing microbiome-based products .....	200
Table 12-1: Companies offering direct-to-consumer genome testing services .....	303

## Figures

Figure 2-1: DNA sequencing process .....	38
Figure 2-2: Components of next generation sequencing.....	39
Figure 2-3: Comparison of traditional sequencing and next generation sequencing .....	47
Figure 2-4: Watson-Crick base pairing.....	48
Figure 2-5: Genome Sequencer FLX system (Roche) .....	52
Figure 2-6: Workflow of Genome Sequencer FLX system .....	53
Figure 2-7: Sequencing by ligation.....	62
Figure 2-8: Construction of SOLiD fragment library using DNA enrichment by ChIP .....	69

Figure 2-9: RNA sequencing .....	70
Figure 2-10: In situ sequencing .....	72
Figure 2-11: Nanopore-based sequence-specific detection of DNA.....	81
Figure 2-12: DNA sequencing through a graphene nanopore .....	82
Figure 2-13: Nanopore-based, single molecule, realtime DNA sequencing .....	90
Figure 2-14: Drop-seq single cell analysis.....	92
Figure 2-15: A scheme of thermosequencing platform .....	96
Figure 2-16: Duplex sequencing .....	100
Figure 3-1: Basic workflow of NGS libraries.....	105
Figure 7-1: Comparison of conventional and high-throughput NGS workflows.....	214
Figure 7-2: Procedure for sequencing of CTC exome .....	220
Figure 7-3: Outline of a study to show current genetic testing and NGS cancer panel .....	239
Figure 8-1: Schematic view of role of WES ND WGS in diagnosis of monogenic disorders.....	248
Figure 10-1: Workflow for pathogen sequencing .....	290
Figure 11-1: Role of sequencing in the development of personalized medicine.....	292