

# **Sequencing**

## **Technologies, Markets & Companies**

by

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

**September 2017**

**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 by training and has been working in the biotechnology/biopharmaceuticals industry for several years. He received graduate training in both Europe and USA, has held academic positions in several countries and is a Fellow of the Faculty of Pharmaceutical Medicine of the Royal College of Physicians of UK. Currently he is a consultant at Jain PharmaBiotech. Prof. Jain's 470 publications include 28 books (5 as editor + 23 as author) and 50 special reports, which have covered important areas in biotechnology, gene therapy and biopharmaceuticals. Recent books include "Textbook of Personalized Medicine" (Springer 2009; Japanese edition 2012; 2<sup>nd</sup> ed Springer 2015), "Handbook of Biomarkers" (Springer 2010, Chinese edition 2016, 2<sup>nd</sup> ed 2017), "Handbook of Nanomedicine" (Springer 2008, Chinese edition 2011, 2<sup>nd</sup> ed 2012, 3rd ed 2017), "Application of Biotechnology in Cardiovascular Disorders" (Springer 2011), "Application of Biotechnology in Neurology" (Springer 2013), and "Application of Biotechnology in Oncology" (Springer 2014). He has edited "Applied Neurogenomics" (Springer 2015).

## **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.

**September 2017**  
**Copyright © 2017 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.

# TABLE OF CONTENTS

<b>0. Executive Summary .....</b>	<b>17</b>
<b>1. Introduction .....</b>	<b>19</b>
<b>Definition and scope of sequencing .....</b>	<b>19</b>
<b>Historical aspects of sequencing.....</b>	<b>19</b>
<b>Basics of molecular biology .....</b>	<b>20</b>
DNA .....	20
<i>DNA polymerases.....</i>	20
<i>Restriction endonucleases.....</i>	21
<i>DNA methylation.....</i>	21
RNA .....	22
<i>RNA polymerases.....</i>	22
<i>Non-coding RNAs.....</i>	22
DNA transcription.....	23
Chromosomes .....	23
<i>Chromatin.....</i>	23
<i>Chromosome sequences .....</i>	24
Telomeres.....	24
Mitochondrial DNA .....	25
Genes.....	25
<i>The genetic code.....</i>	25
<i>Gene expression .....</i>	26
<i>Genes and human diseases .....</i>	26
The human genome .....	27
<i>ENCODE .....</i>	27
The epigenome.....	28
<i>Epigenetics and epigenomics .....</i>	28
<i>DNA methylation.....</i>	28
<i>Human Epigenome Atlas .....</i>	29
Metagenomics .....	30
<b>Variations in the human genome .....</b>	<b>30</b>
Variations in DNA sequences .....	31
Single nucleotide polymorphisms .....	31
Haplotyping.....	31
Complex chromosomal rearrangements .....	32
Insertions and deletions in the human genome .....	32
Large scale variation in human genome .....	33
Variation in copy number in the human genome .....	33
Structural variations in the human genome .....	35
Transposons.....	35
<i>Retrotransposon capture sequencing .....</i>	36
Mapping and sequencing of structural variation from human genomes.....	36
<b>Impact of sequencing on healthcare.....</b>	<b>37</b>
<b>2. DNA Sequencing Technologies.....</b>	<b>39</b>
<b>Introduction .....</b>	<b>39</b>
DNA extraction and sample preparation .....	40
<i>Apollo 300 System for next generation sequencing .....</i>	41
<i>Electrophoresis-based method.....</i>	41
<i>Ion OneTouch System .....</i>	42
<i>Microfluidics-based extraction and sample preparation.....</i>	42
<i>Pressure Cycling Technology .....</i>	42
<i>Selective immobilization of nucleic acids onto magnetic microparticles.....</i>	42
<i>Targeted and hybridization-based DNA capture.....</i>	42
Sanger-sequencing technology .....	43
<i>Dye-terminator sequencing .....</i>	43
<i>Large-scale sequencing.....</i>	44
<i>Automated DNA-sequencing .....</i>	44
<i>Enhancements of Sanger-sequencing .....</i>	44
<i>ABI PRISM® 310 Genetic Analyzer.....</i>	45
<i>Life Technologies' 3500 Dx genetic analyzer.....</i>	45
<b>Limitations of sequencing methods and measures to remedy them .....</b>	<b>45</b>
<i>Paired end transcriptome sequencing to overcome short read lengths .....</i>	46
<i>Long vs short read lengths .....</i>	46
<i>Validation of NGS data.....</i>	47
<b>Emerging sequencing technologies.....</b>	<b>47</b>
<b>Chemical DNA sequencing .....</b>	<b>48</b>

Chemical affinity capture and massively parallel DNA sequencing.....	49
<b>Second generation sequencers .....</b>	<b>50</b>
4300 DNA analyzer .....	50
Apollo 100 .....	50
Applied Biosystems 3500 series Genetic Analyzer .....	50
"Color blind" approach to DNA sequencing .....	51
Cyclic array sequencing .....	51
CEQ™ 8000.....	51
DeepCAGE sequencing.....	52
Electron microscope-based DNA sequencing .....	52
GS-FLEX system (Roche) .....	53
<i>Background of sequencing technology .....</i>	<i>53</i>
IBS sequencing technology .....	57
Illumina Genome Analyzer System .....	58
<i>MiSeqDx .....</i>	<i>59</i>
<i>NextSeq 500 desktop sequencer .....</i>	<i>59</i>
<i>MiSeq FGx.....</i>	<i>59</i>
Ion Torrent's sequencing technology .....	60
<i>Ion S5 system .....</i>	<i>60</i>
MegaBACE 500 .....	61
Microdroplet-based PCR for large-scale targeted sequencing. ....	61
Millikan sequencing .....	61
Multiplex amplification of human DNA sequences .....	62
Nanoscale sequencing .....	62
Polonator sequencer.....	63
RainStorm™ microdroplet technology .....	63
Sequential DEXAS .....	64
SOLiD system: sequencing by ligation .....	64
PCR-based DNA sequencing technologies .....	65
<i>Bridge amplification PCR system .....</i>	<i>66</i>
<i>COLD-PCR and sequencing.....</i>	<i>66</i>
<i>Digital PCR.....</i>	<i>67</i>
<i>Dual primer emulsion PCR.....</i>	<i>67</i>
<i>Emulsion PCR .....</i>	<i>68</i>
<i>Multiplex PCR .....</i>	<i>68</i>
<i>SiMSen-seq for ultrasensitive detection of mutations .....</i>	<i>68</i>
Non-PCR based sequencing .....	69
Nucleic acid sequence-based amplification .....	69
Microarray-based DNA sequencing technologies.....	69
<i>Arrayit's® H25K .....</i>	<i>70</i>
<i>High-throughput array-based resequencing.....</i>	<i>70</i>
<i>Sequencing by hybridization.....</i>	<i>70</i>
<i>SOLiD-System based ChIP-Sequencing .....</i>	<i>71</i>
<i>Companies developing whole genome chips/microarrays.....</i>	<i>72</i>
Next generation sequencing vs microarrays for gene expression profiling.....	72
RNA sequencing.....	73
<i>RNA-seq vs microarrays.....</i>	<i>74</i>
<i>Capture sequencing.....</i>	<i>74</i>
<i>ChIP-Seq .....</i>	<i>74</i>
<i>Strand specific RNA sequencing .....</i>	<i>75</i>
<i>SHAPE-Seq .....</i>	<i>75</i>
<i>Applications of RNA-seq.....</i>	<i>76</i>
<i>Challenges for future development of RNA-seq .....</i>	<i>76</i>
Exome sequencing .....	76
<i>Limitations of exome sequencing .....</i>	<i>77</i>
<i>Human exome microarrays .....</i>	<i>77</i>
<i>WES vs WGS.....</i>	<i>77</i>
Third generation sequencing .....	78
<i>SOLiD4 System.....</i>	<i>78</i>
<i>SOLiD PI System.....</i>	<i>79</i>
<i>Helicos™ Genetic Analysis System .....</i>	<i>79</i>
<i>Molecular Combing.....</i>	<i>80</i>
Nanotechnology-based sequencing.....	80
<i>DNA sequence by use of nanoparticles.....</i>	<i>80</i>
<i>Denaturation mapping of DNA in nanofluidic channels .....</i>	<i>81</i>
<i>Nanopore sequencing .....</i>	<i>81</i>
<i>Sequencing through graphene nanopores.....</i>	<i>84</i>
Convex lens-induced nanoscale templating .....	84
Detection of single molecules for sequencing .....	85
<i>Helicos™ Genetic Analysis System .....</i>	<i>86</i>
<i>Molecular Combing.....</i>	<i>87</i>

<i>Optical Mapping</i> .....	87
<i>Nanopore-based single-molecule detection of specific DNA sequences</i> .....	88
<i>Phasing through the sequencing of single molecules</i> .....	89
<i>Sequencing-by-synthesis for single-molecule sequencing</i> .....	89
<i>Single molecule DNA sequencing by use of carbon nanotubes</i> .....	90
<i>Single molecule sequencing using Qdot nanocrystals</i> .....	90
<i>Single-molecule DNA sequencing in a sTOP chip nanowell</i> .....	91
<i>Single-molecule real-time sequencing</i> .....	91
<i>Single molecule targeted sequencing using GenoCare™ Analyzer</i> .....	93
Single cell sequencing .....	93
<i>MALBAC for single cell DNA sequencing</i> .....	93
<i>Microfluidic single-cell whole-transcriptome sequencing</i> .....	94
<i>Single cell sequencing of uncultured microbes</i> .....	94
<i>Single sperm sequencing</i> .....	94
<i>Single cell RNA sequencing</i> .....	95
<i>Single-cell combinatorial indexed sequencing</i> .....	96
<i>Future of single cell sequencing</i> .....	96
Thermosequencing .....	96
Whole genome sequencing for haplotyping .....	97
Mitochondrial exome sequencing .....	98
ImmunoSEQ technology .....	98
<b>Future prospects of next generation sequencing</b> .....	<b>99</b>
Devices for NGS .....	99
Reduction of errors in NGS .....	99
<i>Artifactual mutations during the sample preparation process</i> .....	99
<i>Contamination in high throughput sequencing</i> .....	100
<i>Duplex sequencing</i> .....	100
<b>3. Role of Bioinformatics in Sequencing</b> .....	<b>103</b>
<b>Introduction</b> .....	<b>103</b>
<b>Growth of the sequencing database</b> .....	<b>103</b>
<b>Sequencing data storage</b> .....	<b>103</b>
Cloud computing for sequencing data .....	104
<i>Cloud computing facilities</i> .....	104
<b>Bioinformatics challenges of new sequencing technology</b> .....	<b>104</b>
<b>Construction of libraries for NGS</b> .....	<b>105</b>
<b>Bioinformatic tools for analysis of genomic sequencing data</b> .....	<b>106</b>
Software for DNA sequencing .....	107
<i>Software from academic and open sources</i> .....	107
<i>Commercial software for sequencing</i> .....	108
ChIA-PET tool for analysis with paired-end tag sequencing .....	110
CLC Cancer Research Workbench .....	110
Compressive genomics .....	110
Detection of CNVs and gene duplications .....	111
Detection of SVs in massively parallel sequencing data .....	111
Differential expression analysis for sequence count data .....	112
DRAGEN™ Bio-IT Processor .....	112
Expression profiling without genome sequence information .....	112
Ingenuity® Variant Analysis™ .....	113
Ion Reporter Software .....	113
Opal platform .....	113
VAAST .....	114
<b>Accessing DNA sequence information</b> .....	<b>114</b>
Clinical Genomicist Workspace for managing NGS-based clinical tests .....	114
<b>Analysis of genomic variation by sequencing of large populations</b> .....	<b>115</b>
Analysis of rare variants in NGS studies .....	115
<b>Human gene connectome</b> .....	<b>116</b>
<b>Funding of research for interpretation of sequencing data</b> .....	<b>116</b>
<b>Future challenges for managing sequencing data</b> .....	<b>117</b>
<b>4. Comparative Analysis of Sequencing Technologies</b> .....	<b>119</b>
<b>General findings of the study</b> .....	<b>119</b>
Sanger versus second generation marketed sequencers .....	119
Common features and differences among second generation sequencers .....	119
<b>Third generation large sequencers</b> .....	<b>120</b>
SOLiD4 versus competing large sequencers .....	120
Illumina's HiSeq sequencer .....	121
<b>Third generation desktop sequencers</b> .....	<b>122</b>
BGI's BGISEQ-500 desk top sequencer .....	122
Illumina's MiSeq sequencer .....	122
Roche GS Junior System .....	123

Life Technologies' Benchtop Ion Proton™ Sequencer .....	123
Oxford Nanopore's MinION™ system .....	124
<b>The ideal desk-top sequencer .....</b>	<b>124</b>
<b>SWOT analysis of small sequencers .....</b>	<b>124</b>
<b>Concluding remarks on SWOT analysis of desktop sequencers .....</b>	<b>128</b>

## **5. Sequencing for Research ..... 131**

### **Introduction ..... 131**

#### **Applications in basic research..... 132**

Barcoding of synapses in the brain .....	132
ChIA-PET technology for 3D study of the genome .....	132
ChIP-Seq for study of gene expression .....	133
Chromatin profiling by direct DNA sequencing .....	133
Discovery of immunoglobulin gene by pyrosequencing.....	133
Epigenetic modifications analyzed by next generation sequencing .....	134
Exome sequencing for study of human variation .....	134
Genome sequencing with combinatorial probe anchor ligation.....	135
Identifying protein-coding genes in genomic sequences .....	135
Mutation rate measured by direct sequencing .....	136
Protein-protein interactome network mapping .....	137
RNA sequencing.....	137
<i>Applications of RNA sequencing .....</i>	<i>137</i>
<i>Contribution of RNA structure to gene regulation revealed by RNA-Seq.....</i>	<i>137</i>
<i>Molecular indexing for quantitative targeted RNA sequencing .....</i>	<i>138</i>
Sequencing for the study of microchimerism .....	138
Sequencing for the study of CNVs .....	139
<i>CNVnator .....</i>	<i>139</i>
<i>GS-FLX sequencing for simultaneous detection of mutations and CNVs .....</i>	<i>140</i>
Sequencing the transcriptomes of stem cells .....	140
<i>Sequencing, stem cells and neurodegeneration.....</i>	<i>140</i>
<i>Sequencing, stem cells and regeneration .....</i>	<i>141</i>
Sequencing and synthetic biology .....	141
<i>Synthetic sequence in a bacterial cell .....</i>	<i>141</i>
<i>Synthesizing long DNA molecules.....</i>	<i>142</i>
<i>Functional synthetic proteins .....</i>	<i>142</i>
Single-cell sequencing for study of cell lineage .....	143

#### **Sequencing of human genomes ..... 143**

100,000 Genomes Project .....	143
Genome of the Netherlands.....	144
Missing human genome sequences.....	145
Personal genome sequencing.....	145
Role of sequencing in identification of human remains .....	146
Saudi Arabian human genome program .....	146
Sequence map of the human pan-genome .....	146
Sequencing of African genomes .....	147
Sequencing of Korean genomes .....	147
Sequencing mitochondrial genome.....	148
Sequencing of ancient genomes.....	148
<i>Hominin genome.....</i>	<i>149</i>
<i>Neandertal genome.....</i>	<i>149</i>
<i>Anzick-1 genome .....</i>	<i>150</i>
<i>Genome of the Kennewick Man .....</i>	<i>151</i>
<i>Saqqaq genome.....</i>	<i>151</i>
<i>Sequencing and controversies about origin of Native Americans.....</i>	<i>152</i>
Whole genome sequencing.....	152
Whole-genome sequencing of methylome .....	153
Whole genome resequencing .....	153
Prospects of genome sequencing for human ancestry .....	153

#### **Sequencing genomes of non-human primates ..... 154**

Sequencing of chimpanzee genome.....	154
Sequencing of macaque genome.....	154
Sequencing of gorilla genome.....	155

#### **Sequence the genomes of vertebrate species ..... 155**

#### **Sequencing genomes of other organisms ..... 156**

Species biodiversity study by the Canadian Centre for DNA Barcoding .....	157
Ant genome .....	157
Avian genomes .....	158
Bat genome .....	158
Body louse genome.....	159
Camel genome .....	159
Cat genome .....	159

Dog genome.....	160
Frog genome.....	161
Goat genome .....	162
Horse genome .....	163
<i>Ancient horse genome</i> .....	163
House fly .....	164
Mosquito genome.....	164
Mouse genome .....	164
Tibetan antelope genome.....	165
Turkey genome .....	165
Water flea genome.....	166
Whale genome .....	166
Woolly mammoth genome .....	167
<b>Sequencing studies of the human microbiome.....</b>	<b>167</b>
Human Microbiome Project.....	168
Human virome.....	168
Microbial codes to distinguish individuals .....	169
<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>185</b>
<b>Introduction .....</b>	<b>185</b>
<b>Applications of sequencing in molecular diagnostics .....</b>	<b>185</b>
Clinical exome sequencing .....	186
<i>ACE Clinical Exome test</i> .....	186
Diagnosis and screening of genetic disorders.....	187
<i>Cystic fibrosis screening using high-throughput NGS</i> .....	187
<i>CNV sequencing for diagnosis of chromosomal disorders</i> .....	187
<i>Genetic misdiagnosis due to limited population sequencing data</i> .....	188
<i>Karyotyping based on NGS</i> .....	188
<i>NGS for diagnosis of rare genetic disorders</i> .....	188
<i>Role of WGS in screening of newborns</i> .....	189
<i>Role of WGS in chromothripsis</i> .....	189
<i>Single cell sequencing for PGD</i> .....	190
<i>WGS for identification of genetic disorders in critically ill infants</i> .....	190
<i>WGS for pre-implantation genetic diagnosis in IVF</i> .....	190
Guidelines for use of sequencing for diagnosis .....	191
Incidental findings on clinical sequencing .....	191
NGS for diagnosis of CNS infections .....	192
NGS for detection of solid organ transplant rejection .....	192
NGS for forensic diagnosis .....	193
NGS-based molecular diagnosis at POC using mobile phone microscopy.....	193
Companies developing sequence-based molecular diagnostics .....	194
<b>Sequencing in cardiovascular disorders.....</b>	<b>195</b>
Inherited cardiomyopathies.....	195
Exome sequencing and mutations associated with risk of coronary heart disease .....	196
<b>Sequencing for study of the human immune system.....</b>	<b>196</b>
Sequencing for investigating drug-virome interactions in organ transplants.....	197

Immune profiling using a synthetic human virome .....	197
NGS-based HLA typing .....	198
<i>High-throughput HLA genotyping with deep sequencing</i> .....	198
<i>Approaches to NGS-based HLA typing</i> .....	199
<i>Applications of NGS-based HLA typing</i> .....	200
<b>Applications of sequencing relevant to the human microbiome .....</b>	<b>200</b>
Sequencing the gut microbiome for discovery of immunomodulators .....	200
Sequencing of gut microbes in obesity .....	201
Companies developing microbiome-based products .....	201
<b>Pharmaceutical applications of sequencing .....</b>	<b>202</b>
Drug discovery and development .....	202
Resequencing .....	203
RNA profiling .....	203
Quantitative selection of aptamers through sequencing .....	204
Next generation sequencing and drug safety .....	204
<b>Sequencing in aging research .....</b>	<b>205</b>
WGS for the study of supercentenarians .....	205
<i>Sequencing of APOE gene in centenarians</i> .....	205

## **7. Applications of Sequencing in Oncology .....** **207**

<b>Introduction .....</b>	<b>207</b>
<b>Sequencing technologies for cancer .....</b>	<b>207</b>
A project to assess sequencing technologies for tumor DNA .....	207
A universal NGS-based oncology test system .....	208
Amplicon sequencing in cancer .....	208
ASCO guidelines for use of NGS to test cancer susceptibility .....	208
Cancer Genome Atlas .....	209
Catalog of cancer genes .....	209
Detection of cancer biomarkers .....	210
<i>Sequencing mitochondrial DNA to identify cancer biomarkers</i> .....	210
<i>Biomarkers for personalizing cancer treatment</i> .....	211
Digital proteomics for cancer profiling .....	212
Discrepancies in cancer genome sequencing .....	212
Epigenome profiling .....	212
Exosome sequencing .....	213
Gaining insights into mutational processes .....	213
Multiple-gene sequencing panel to assess risk of hereditary cancer .....	213
Multiplexed cancer gene mutation analysis .....	214
NGS for enhancing verification rate of chromosomal structural rearrangements .....	214
NGS-based molecular profiling of cancer in FFPE specimens .....	216
Paired-end sequencing .....	217
Pathology tissue-ChIP .....	218
Quality control of NGS in oncology .....	218
RNA-Seq to study cancer transcriptome .....	218
Sequencing cancer cell lines .....	219
<b>Sequencing for studying somatic mutations in cancer .....</b>	<b>219</b>
Chromothripsis in cancer .....	219
Microsatellite instability in cancer genome .....	220
Somatically acquired genomic rearrangements in cancer .....	220
Sequencing for identification of FGFR gene fusions in cancer .....	220
Sequencing single cells to study evolution of cancer .....	221
Sequencing circulating tumor cell genomes .....	221
<b>High throughput sequencing for anticancer drug discovery .....</b>	<b>222</b>
<b>NGS for developing targeted cancer therapies .....</b>	<b>223</b>
<b>Sequencing for assessing resistance to anticancer therapy .....</b>	<b>223</b>
<b>Sequencing of tumors of various organs .....</b>	<b>223</b>
Brain tumors .....	223
<i>Sequencing for genetic alterations in gliomas</i> .....	224
<i>Sequencing for genetic alterations in medulloblastoma</i> .....	224
Breast cancer .....	224
<i>BRCA mutations</i> .....	225
<i>Circulating nucleic acids as biomarkers of cancer</i> .....	225
<i>Deep sequencing of miRNA for signatures of invasiveness</i> .....	226
<i>NGS reveals heterogeneity of breast cancer</i> .....	226
<i>Sequencing of breast cancer metastases</i> .....	226
<i>Triple negative breast cancer</i> .....	227
<i>Whole genome sequencing in breast cancer</i> .....	227
Colorectal cancer .....	228
Gastric cancer .....	229
Head and neck cancer .....	229
<i>NGS for detection of HPV sequences in carcinoma of oropharynx</i> .....	229



Hematological malignancies .....	230
<i>Acute myeloid leukemia</i> .....	230
<i>Acute promyelocytic leukemia</i> .....	231
<i>Chronic myelomonocytic leukemia</i> .....	231
<i>Hairy-cell leukemia</i> .....	232
<i>Hematological cancer risk inferred from blood DNA sequence</i> .....	232
<i>Myelodysplastic syndromes</i> .....	232
<i>Sequencing in chronic neutrophilic leukemia and atypical CML</i> .....	233
Sequencing in hepatocellular carcinoma .....	233
Lung cancer .....	234
<i>Sequencing of small cell lung cancer</i> .....	234
Melanoma .....	234
Ovarian cancer .....	235
Prostate cancer .....	235
<i>Identification of mutations in prostate cancer by exome sequencing</i> .....	236
<i>Role of sequencing in liquid biopsy for prostate cancer patients</i> .....	237
Current status and future prospects of NGS applications in oncology .....	237
<i>Actionable Genome Consortium to guide NGS in clinical oncology</i> .....	238
<i>Points to consider for germline findings in tumor-only sequencing</i> .....	239
<i>Translation of gene panel sequencing into routine cancer diagnosis</i> .....	239
<b>8. Sequencing in Genetic Disorders .....</b>	<b>241</b>
<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 .....	246
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>251</b>
Bartter syndrome .....	251
CHARGE syndrome .....	251
DiGeorge syndrome .....	251
Discovery of the gene for Miller syndrome .....	252
Discovery of the gene for Kabuki syndrome .....	252
Familial combined hypolipidemia .....	253
Familial thoracic aortic aneurysm .....	253
Hereditary ataxia .....	253
Hereditary blindness .....	253
Neurofibromatosis type 1 .....	254
Noonan syndrome .....	255
Proteus syndrome .....	255
Syndrome of hypogonadotropic hypogonadism, ataxia, and dementia .....	256
Syndrome of polyarthritis nodosa vasculopathy .....	256
X-linked disorder due to N-terminal acetyltransferase deficiency .....	256
Sequencing in mitochondrial disorders .....	257
<i>Chronic progressive external ophthalmoplegia</i> .....	257
Role of NGS in prevention of chromosome abnormalities & monogenic disorders .....	258
<b>9. Sequencing in Neurological and Psychiatric Disorders .....</b>	<b>259</b>
<b>Introduction .....</b>	<b>259</b>
<b>Sequencing in Alzheimer disease .....</b>	<b>259</b>
<b>Sequencing in Parkinson disease .....</b>	<b>261</b>
<b>Sequencing in Huntington's disease .....</b>	<b>262</b>
<b>Sequencing in Wilson's disease .....</b>	<b>262</b>
<b>Sequencing in ataxias .....</b>	<b>263</b>
<b>Sequencing in epilepsy .....</b>	<b>263</b>
Epileptic encephalopathy .....	263
<b>Sequencing for mutations in familial amyotrophic lateral sclerosis .....</b>	<b>264</b>
<b>Sequencing of whole genome in Charcot-Marie-Tooth disease .....</b>	<b>264</b>
<b>Sequence-based detection of a variant of Lambert-Eaton syndrome .....</b>	<b>265</b>
<b>Sequencing in muscular dystrophy .....</b>	<b>265</b>
<b>Sequencing in acute brain injury due to hemorrhage .....</b>	<b>266</b>
<b>Sequencing for mutations in autism spectrum disorders .....</b>	<b>266</b>
<b>Sequencing for diagnosis of intellectual disability .....</b>	<b>267</b>

<b>Sequencing in neurodevelopmental disorders</b> .....	<b>267</b>
NGS for identifying mutations in RNA gene.....	267
<b>Sequencing in Möbius syndrome</b> .....	<b>268</b>
<b>Sequencing in attention-deficit/hyperactivity disorder</b> .....	<b>268</b>
<b>Sequencing in schizophrenia and bipolar disorder</b> .....	<b>268</b>
<b>Sequencing in drug addiction</b> .....	<b>268</b>
<b>10. Applications of sequencings in infections</b> .....	<b>271</b>
<b>Introduction</b> .....	<b>271</b>
<b>DNA sequencing for study of bacterial epidemics</b> .....	<b>271</b>
Sequencing of GAS genotype emm89 .....	272
Genome sequencing of H. influenzae to identify population structure.....	272
Role of sequencing in cholera epidemics .....	272
Role of sequencing in epidemic of Shiga toxin-producing E. coli .....	273
Sequencing study of emergence of M. tuberculosis in East Africa .....	274
Sequencing study of Salmonella emergence in Sub-Saharan Africa .....	274
Sequencing of ancient specimes from past epidemics .....	275
<b>Sequencing for tracking hospital acquired infections</b> .....	<b>275</b>
Sequencing for investigation of MRSA outbreaks .....	275
Role of sequencing in tracking a hospital infection of K. pneumoniae .....	276
Role of whole genome sequencing in identification of C. difficile .....	277
Role of NGS in diagnosis of infectious agents causing meningitis & encephalitis .....	277
<b>Role of sequencing in the management of bacterial infections</b> .....	<b>277</b>
Pyrosequencing of microbial flora in leg ulcers .....	277
Sequencing for mapping genomic variation in Mycobacterium ulcerans .....	278
<b>Sequencing in the management of antimicrobial drug resistance</b> .....	<b>278</b>
Maximum-depth sequencing.....	278
Sequencing for study of antibiotic resistance in bacteria .....	279
Sequencing for predicting the virulence of MRSA.....	279
Sequencing for detection of drug resistance in Plasmodium falciparum .....	279
WGS for diagnosis of drug-resistant M. tuberculosis .....	280
<b>Sequencing for investigation of food-borne infections</b> .....	<b>280</b>
<b>Sequencing for mapping genetic interactions in bacteria</b> .....	<b>280</b>
<b>Metagenomic sequencing of bacteria</b> .....	<b>281</b>
Sequencing of DNA from single cells of bacteria .....	281
<b>Sequencing of the fungal genomes</b> .....	<b>282</b>
<b>Sequencing of human salivary microbiome</b> .....	<b>282</b>
<b>Next generation sequencing for antibacterial therapeutic discovery</b> .....	<b>283</b>
<b>High throughput sequencing for diagnosis of viral diseases</b> .....	<b>283</b>
<b>Sequencing in the management of HIV/AIDS</b> .....	<b>284</b>
Long read sequencing for personalizing HIV therapy.....	284
NGS for studying neuroAIDS .....	284
Sequencing plus immunological analyses to study HIV evolution .....	285
Surveillance of drug resistance in HIV-infected individuals .....	285
<b>Sequencing in the management of Ebola virus infection</b> .....	<b>286</b>
<b>Sequencing in the management of HCV</b> .....	<b>286</b>
Sequencing for detection of a novel pegivirus associated with HCV.....	287
<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 .....	288
Surveillance of H1N1 influenza A virus using resequencing arrays .....	288
<b>Regulatory aspects of sequencing for diagnosis of infections</b> .....	<b>288</b>
<b>11. Role of Sequencing in Personalized Medicine</b> .....	<b>291</b>
<b>Introduction</b> .....	<b>291</b>
<b>Technologies relevant to sequencing and personalized medicine</b> .....	<b>291</b>
Whole genome sequencing and personalized medicine.....	291
Whole exome sequencing and personalized disease risk.....	292
Large-scale deep sequencing of human genomes .....	293
<b>Personal Genome Project</b> .....	<b>293</b>
<b>Role of sequencing in personalized cancer management</b> .....	<b>294</b>
<b>Standardization of sequencing for personalized medicine</b> .....	<b>295</b>
Regulating genomic testing in the era of personalized medicine .....	295
<b>Future of sequencing and personalized medicine</b> .....	<b>296</b>
<b>12. Current Status and Future Prospects</b> .....	<b>299</b>
<b>ACMG clinical laboratory standards for NGS</b> .....	<b>299</b>
<b>Regulatory and quality control issues of sequencers</b> .....	<b>299</b>
<b>Applications of NGS in clinical trials</b> .....	<b>300</b>
Human transcriptome array in clinical trials.....	300
<b>Challenges for clinical applications of NGS</b> .....	<b>301</b>

Direct-to-consumer WGS services .....	302
<b>Ethical aspects of sequencing .....</b>	<b>303</b>
<b>Future trends in clinical sequencing .....</b>	<b>303</b>
Rare Diseases Genomes Project.....	303
<b>13.    Markets for Sequencers .....</b>	<b>305</b>
<b>Introduction .....</b>	<b>305</b>
<b>Methods used for estimation of sequencer markets.....</b>	<b>305</b>
<b>Currently marketed sequencers .....</b>	<b>305</b>
<b>Academic and research markets for sequencing .....</b>	<b>306</b>
<b>Factors affecting future development of sequencing markets.....</b>	<b>306</b>
Future needs and support of research.....	306
Bioinformatics in relation to sequencing .....	307
Cost of integrating WGS into clinical care .....	307
Reducing the cost of human genome sequencing .....	307
<i>US Government-supported research on sequencing.....</i>	<i>308</i>
<i>Contribution of American Recovery and Reinvestment Act.....</i>	<i>310</i>
<i>Genome X Prize Foundation.....</i>	<i>310</i>
<i>Innovations to reduce cost of whole genome sequencing .....</i>	<i>311</i>
<i>Commercial aspects of low cost genome sequencing .....</i>	<i>311</i>
<i>Genome sequencing suitable for personalized medicine .....</i>	<i>312</i>
<b>Global sequencing markets.....</b>	<b>312</b>
Global markets for sequencers.....	312
Markets for sequencing services according to geographical regions .....	313
Global sequencing markets according to applications.....	313
Global sequencing markets according to therapeutic areas.....	314
<b>Market trends for NGS .....</b>	<b>314</b>
Needs of the clinical market for NGS .....	315
Sequencers for the clinical market .....	315
<b>Challenges to developing market for sequencers.....</b>	<b>316</b>
<b>Recommendations .....</b>	<b>317</b>
<b>14.    Companies Involved in Sequencing .....</b>	<b>319</b>
<b>Introduction .....</b>	<b>319</b>
<b>Major players in sequencing .....</b>	<b>322</b>
<b>Profiles of companies involved in sequencing .....</b>	<b>323</b>
<b>Collaborations.....</b>	<b>486</b>
<b>15.    References.....</b>	<b>493</b>

## Tables

Table 1-1: Historical landmarks in DNA sequencing .....	20
Table 1-2: Genetic variations in the human genome.....	30
Table 2-1: ChIP detection platforms for sequencing .....	71
Table 2-2: Companies developing whole genome chips/microarrays.....	72
Table 2-3: Systems for single molecule sequencing.....	85
Table 3-1: Software programs for sequencing from open sources .....	107
Table 3-2: Companies providing DNA sequencing software .....	108
Table 4-1: Comparison of a generation I and generation II sequencers.....	119
Table 4-2: Similarities and differences between second generation sequencers.....	120
Table 4-3: SWOT of ABI 310 .....	125
Table 4-4: SWOT of IBS sequencing.....	125
Table 4-5: SWOT of NABsys' Hybridization-Assisted Nanopore Sequencing.....	125
Table 4-6: SWOT of 4300 DNA Analysis System Li-Cor.....	125
Table 4-7: SWOT of BGI's BGISEQ-500 desktop sequencer .....	125
Table 4-8: SWOT of Polonator .....	126
Table 4-9: SWOT of Roche GS FLEX Junior.....	126
Table 4-10: SWOT of Oxford Nanopore's MinION™ system.....	126
Table 4-11: SWOT of Ion Torrent™ Personal Genome Machine .....	126
Table 4-12: SWOT of Pacific BioSciences' single-molecule real-time sequencing.....	127
Table 4-13: SWOT of Illumina's MiSeqDx.....	127
Table 4-14: SWOT of Illumina's NextSeq 500 .....	127
Table 4-15: SWOT of QIAGEN's GeneReader™ sequencer .....	127
Table 4-16: SWOT of Bio-Rad (formerly GnuBio's) droplet-based sequencing system .....	128
Table 4-17: SWOT of 10X Genomics' GemCode .....	128
Table 5-1: Number of genes in organisms with fully sequenced genomes.....	131
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 .....	194
Table 6-2: Companies developing microbiome-based products .....	201
Table 12-1: Companies offering direct-to-consumer genome testing services .....	302
Table 13-1: Marketed next generation sequencers .....	305
Table 13-2: De novo sequencing vs resequencing markets.....	306
Table 13-3: Global markets for sequencers from 2016 to 2026.....	313
Table 13-4: Global markets for sequencing services according to geographical regions.....	313
Table 13-5: Global markets for sequencing services according to applications .....	313
Table 13-6: Sequencing markets according to therapeutic areas from 2016 to 2026 .....	314
Table 14-1: Companies developing sequencing technologies and instruments .....	319
Table 14-2: Companies that provide sequencing services.....	320
Table 14-3: Companies that provide bioinformatics support for sequencing.....	322
Table 14-4: Major players in sequencing.....	322
Table 14-5: Selected collaborations for DNA sequencing .....	486

## Figures

Figure 2-1: DNA sequencing process .....	39
Figure 2-2: Components of next generation sequencing.....	40
Figure 2-3: Comparison of traditional sequencing and next generation sequencing .....	48
Figure 2-4: Watson-Crick base pairing.....	49
Figure 2-5: Genome Sequencer FLX system (Roche) .....	55
Figure 2-6: Workflow of Genome Sequencer FLX system .....	56
Figure 2-7: Sequencing by ligation.....	65
Figure 2-8: Construction of SOLiD fragment library using DNA enrichment by ChIP.....	71
Figure 2-9: RNA sequencing .....	73
Figure 2-10: Nanopore-based sequence-specific detection of DNA.....	83
Figure 2-11: DNA sequencing through a graphene nanopore .....	84
Figure 2-12: Nanopore-based, single molecule, realtime DNA sequencing .....	92
Figure 2-13: A scheme of thermosequencing platform .....	97
Figure 2-14: Duplex sequencing .....	100
Figure 3-1: Basic workflow of NGS libraries.....	106
Figure 7-1: Comparison of conventional and high-throughput NGS workflows.....	216
Figure 7-2: Procedure for sequencing of CTC exome .....	222
Figure 7-3: Outline of a study to show current genetic testing and NGS cancer panel .....	240
Figure 8-1: Schematic view of role of WES ND WGS in diagnosis of monogenic disorders.....	249
Figure 11-1: Role of sequencing in the development of personalized medicine.....	291
Figure 13-1: Cost of sequencing per genome .....	310
Figure 13-2: Global markets for sequencing services according to applications .....	314