

Index

- 10 element, 92, 131–133
- 35 element, 92–93, 131–133, 153
- 5′-RACE, 124, 129
- α- proteobacteria, 20
- A+T content, 9, 133, 136–137, 177–178, 181, 185
- Acinetobacter baumannii*, 74, 87
- activation, 96, 146, 150, 153, 156, 162–165, 170, 174
- adapter sequences, 62, 81
- alarmone, 97, 186
- alignment of short reads, 70
- antibiotic resistance, 53, 56, 75
- antigen, 26, 49, 51, 87
- antigenic variation, 26–27, 56
- anti-sense RNA, 99
- antisense transcription, 125
- Bacillus anthracis*, 122
- background subtraction, 39–40, 44–45
- bacteriophage, 4–5, 29, 53, 56, 193
- basecomposition profile, 10
- Baumannia cicadellincola*, 24–25
- bioconductor, 39, 41, 43–44, 112
- bipartite promoter, 132–134
- bisulphite sequencing, 195–197
- Black Death, 75–78
- BLAST, 5, 14, 54, 70, 155
- bowtie, 72
- Buchnera* sp., 23
- Burkholderia dolosa*, 86
- Burrows–Wheeler Transformation, 70–71
- BWA, 72, 76
- BWsearch, 72
- Campylobacter jejuni*, 26
- Candidatus Hodgkinia cicadicola*, 28
- Caulobacter crescentus*, 141–142, 203
- CDD, 14, 155
- cDNA arrays, 45, 54
- chemostat, 85
- ChIP (Chromatin Immuno Precipitation), 118
 - ChIP-chip, 37, 117–119, 129, 145–148, 151–152, 165, 172–175, 179–180, 182, 184–185, 188
 - ChIP-seq, 117–123, 165, 172, 175–178, 182, 185–186, 189
- chromosome, 6, 10, 15, 29, 46–47, 53, 68, 83, 90, 94, 101, 103, 109, 134, 137–142
- Chromosome Conformation Capture, 139–140
- clonal amplification, 62
- clonally amplified templates, 62
- closed complex, 91
- CLR (context likelihood of relatedness), 168
- clustalW, 51, 77
- clustering analysis, 50, 55, 58, 179
- coding density, 24, 29
- coding potential, 22, 24
- COG, 31
- combinatorial control, 159–160, 168
- commensal, 20
- comparative genome hybridisation, 20, 36, 45, 47, 52, 56, 78, 100, 113, 138
- comparative genomics, 4–6, 15, 31, 33, 192
- constrained supercoiling, 93
- contact frequency or contact probability, 139

208 Index

- contig, 7–8, 68, 74, 76, 88, 105, 109
- contingency loci, 26–27, 56, 191
- core variable, 53, 55
- co-regulatory networks, 160
- cross-regulation, 169
- CRP, 159–160, 164, 167–168, 172–175, 177
- cufflinks, 107
- cyclic reversible termination, 64–65

- deamination, 76
- de Bruijn graphs, 67, 69, 108
- deep sequencing, 2–3, 60, 77, 79, 81, 102–103, 119, 122–123, 127, 129, 139, 141, 186, 189, 195, 198
- de-novo* transcriptome assembly, 108–109
- depth of coverage, 2, 63, 100, 105–106, 108–110, 124
- DEseq, 112–114, 121
- dibase encoding, 66
- differential RNA-seq (dRNA-seq), 127
- DNA compaction, 137
- DNA gyrase, 86, 93, 134–135, 179
- DNA ligase, 64
- DNA looping, 158
- DNA methylation, 191–193, 195, 200–204
- DNA methyltransferase, 192–194, 202–203
- DNA microarrays, 36, 38, 60, 89, 99, 101, 114, 139
- DNA polymerase, 64–65, 83, 101, 125, 200
- DNA relaxation, 134–135, 137
- DNA repair, 27, 151
- DNA supercoiling, 93–94, 97, 134–137, 144, 179
- DNA topology, 134–135
- draft assembly, 74, 83
- draft genome, 76, 79, 100

- EasyGene, 51
- EdgeR, 112–114, 121
- EggNOG, 31
- emulsion PCR, 62–63
- ENCODE, 70, 125
- endosymbiont, 20, 23–24, 28–29
- epigenetic, 2, 191, 195, 203, 205
- Escherichia coli*, 1, 6, 8–9, 16, 46–48, 79, 90, 128, 136, 155, 167
 - E. coli* K12, 9, 11–12, 16–17, 30, 33, 49, 70, 95
- Eulerian path, 67–68
- evolution, 1, 4, 19–20, 27–28, 48, 53, 84–85, 86–87, 117, 134, 171, 180, 194
- excisive recombination, 22
- experimental annotation, 2, 37, 114, 125, 146, 149, 188, 205
- experimental genome annotation, 99–100, 102, 106, 122–123, 128, 146
- exponential distribution, 40–41
- Expressed Sequence Tag, 99
- extended -10 element, 131–132

- FIS, 160, 174–177, 180, 183, 185
- fluorescence, 37–41, 43–45, 62, 64, 100, 115–116, 143, 179–180, 200
- fluorescent reporter, 100, 114–116
- FNR, 160, 185
- frameshifting, 26, 72
- functional categories, 31, 33, 161

- G/C skews, 9
- G+C content, 9–10, 16, 110, 124
- GACK, 45, 49–50
- gap closure, 6–7
- gene annotation, 6, 49, 51
- gene decay, 22–23, 33
- gene degradation, 22
- gene duplication, 29, 151
- gene expression microarray, 37, 126, 151, 175, 180, 182, 186
- gene expression profile, 96, 99–100, 166–168, 172
- gene finding, 6, 10
- gene loss, 21–22, 29, 47, 59
- gene ontology, 14
- gene order, 15, 22–23, 30
- gene regulatory mechanism, 1, 191
- gene regulatory network, 2, 130, 168
- gene relic, 21
- GeneMark, 12
- general stress response, 152, 187
- genetic content, 1–2, 80, 89, 191
- genetic diversity/genetic variation, 1, 20, 36, 56, 60
- genome annotation, 99–100, 102, 106, 122–124, 128–129, 146
- genome assembly, 7, 16, 62, 67, 69, 74, 81, 105, 108–109

Cambridge University Press

978-1-107-07983-0 - Bacterial Genomics: Genome Organization and Gene Expression Tools

Aswin Sai Narain Seshasayee

Index

[More information](#)

- genome reduction, 1, 19, 22, 24, 27, 29
- genome sequencing, 1–2, 6–8, 16, 24, 27, 32, 36, 56, 59, 61, 66, 73–74, 82–83, 105
- genome size, 18, 20, 28–29, 33–34, 51, 53, 154, 203
- genome streamlining, 23, 29
- genome-tiling microarrays, 99
- genomic epidemiology, 36
- genomic variability, 20
- genomics, 1–6, 15, 25, 31, 33, 80–81, 192, 195
- Geobacter sulfurreducens*, 149–150
- GLIMMER, 12
- GO, 31, 33

- Haemophilus*, 5, 26–27
- Hamiltonian path, 66, 68
- Helicobacter pylori*, 26, 46, 52, 56–57, 191, 193
- Heliscope, 61–62, 64
- helix-turn-helix (HTH), 154
- heterochromatin, 180
- Hfq, 99, 128, 188–190
- Hi-C, 140, 144–145
- HMMER, 155
- H-NS, 160, 176, 180–186
- homology, 4, 10, 13–15, 22, 57, 156
- homopolymeric repeats, 26
- homopolymeric tracts, 75
- horizontal gene transfer, 1, 27, 29, 53, 180, 194
- HU, 177–179
- Human Microbiome Project, 73
- hybridisation, 1, 20, 36–37, 39, 43, 45, 47, 49–52, 54, 56–57, 60, 66, 78, 99–100, 113, 119, 138, 188

- IHF, 160, 176–178, 180, 183, 185
- illumina, 60–62, 64, 66, 70, 72–74, 76, 81, 84–86, 104, 123, 197–198
- indels, 70, 88
- INFERNAL, 13
- intron, 107
- in-vitro evolution, 134

- Justinian plague, 75, 78
- laboratory evolution, 84

- lagging strand, 9, 29, 125
- last common ancestor, 86–87
- leaderless mRNA, 128
- leading strand, 17, 29, 125
- library size, 110–112
- limma, 44
- linear regression, 42, 199
- linking number, 93
- local regression, 42
- local TF, 160–164, 173
- LOESS, 42
- looped domain architecture, 138
- low complexity regions, 81
- LRP, 160, 168, 172–173

- M-A plot, 42
- MAQ, 70
- Markov model, 10–12
- mate-pair sequencing, 68, 69
- MeDIP, 195, 198–200
- messenger RNA or mRNA, 90
- meta-analysis, 28
- metagenomic, 80–82
- microarrays, 2, 36–38, 45–46, 52, 54, 60, 78, 89, 99–102, 113–115, 118–119, 122, 126–127, 129, 134, 138–139, 151, 165, 168, 172, 174–175, 180, 182, 186–187, 195
- mismatch probe, 39–41, 45
- mitochondria, 20, 22, 29
- mixed distribution, 41
- mobile genetic elements, 53–55, 86
- modENCODE, 70, 125
- morbidity, 85
- mRNA splicing, 105
- mRNA stability, 98
- MRSA, 52, 54
- multi-gene families, 23, 29
- multi-locus sequence typing, 55
- multiple displacement amplification, 83
- multiple sequence alignment, 51, 131, 155, 171
- Mycobacterium tuberculosis*, 23, 128
- Mycobacterium leprae*, 1, 22–23, 29
- Mycoplasma pneumoniae*, 26

- nanopore, 61
- negative binomial distribution, 113–114

210 Index

- negative supercoiling, 93, 134, 137
Neisseria, 26–27
 network motif, 162–164, 168
 next-generation sequencing, 60–61, 70, 73–74, 78, 81, 86, 101
 non-coding RNA, 90–91, 116, 122, 128, 148
 normal distribution, 40–41
 normexp, 41
 nucleoid-associated protein, 160, 173–174, 176
 nylon membrane, 47
 Oases, 108
 open complex, 91, 132, 147
 operon, 9, 13, 17, 30, 97, 123–124, 126–129, 146, 149, 156, 159, 162, 164–165, 168–169, 171, 174, 188
 paired-end sequencing, 7–8, 68–69, 74, 106, 120, 140, 143–144
 pan-genome, 20, 50–51
 pathogenicity island, 53, 56–58, 181, 190
 pathovar, 20, 49, 51
 PCR, 8, 52, 54, 56–57, 62–63, 74, 83, 106, 110, 118, 124, 141, 195, 197, 199
 peak, 28, 120–121, 146, 177
 peak callers, 120
Pelagibacter ubique, 24
 perfect match probe, 39–40
 PFAM, 14, 155
 phages, 24, 32, 55
 phylogenetic analysis, 2, 21, 77, 86–87
 phylogenetic tree, 47, 77–78, 87
 phylogenomics, 36, 45
 plasmid, 55, 116, 135, 184–185, 190, 193, 203
 plasticity zone, 57–58
 Poised RPO, 146, 147
 Poisson distribution, 6, 113–114
 polymerase slippage, 26
 population bottleneck, 21, 29
Porphyromonas gingivalis, 83
 positive supercoiling, 93
 power law, 33
 probe, 1–2, 36–41, 44–45, 50–52, 55–56, 65–66, 116, 134–135, 137, 148, 173
 probiotic, 48, 51
 promoter, 84–85, 90–93, 95–98, 100, 115–117, 122, 128, 131–134, 145–148, 152–153, 156–158, 164–165, 173–174
 promoter activity, 85, 115–116, 164
 promoter capture, 84, 134
 protein structure annotation, 14
 protein-coding, 10, 12–13, 16–17, 22–23, 29, 91, 126, 149, 199
Proteus mirabilis, 20, 30
 pseudogene, 21, 23–24, 29
 pyrosequencing, 61, 64–65, 74–75, 77, 80, 106
 quantile normalisation, 43, 113
 read length, 61, 80
 reference genome, 45–46, 49–50, 53–54, 56, 70, 72, 76, 84, 86, 100, 104, 106–109, 120–121, 123, 144, 195–197
 reference-assisted assembly, 106–107, 109
 regulatory RNA, 12–13, 98–99
 regulonDB, 92, 95, 117, 130–133, 154, 156–159, 162, 168, 171, 173, 177
 RelA, 187
 repeat sequence, 22, 26
 repetitive DNA, 24
 replication, 9–10, 17, 29–31, 33, 49, 83, 93, 101–103, 125, 137, 141–142, 153, 182, 187, 193, 199, 204
 replication fork, 24, 30, 101
 repression, 96–97, 156, 163
 restriction-modification system, 57, 59, 192, 194
 RFAM, 12–13
 Rho, 99, 185–186
Rhodobacter sphaeroides, 151
 Ribosomal RNA, 97, 123
 riboswitches, 12–13, 98–99
Rickettsia conorii, 22
Rickettsia prowazekii, 21
 RIP-ChIP, 188–190
 RMA, 41, 43
 RMAP, 70
 RNA, 2, 4, 12–13, 16, 62, 72, 90, 91, 93, 96–100, 102–107, 111, 114, 116–117, 121–123, 126–129, 133, 135, 145, 148, 153, 179, 181, 187–190
 RNA polymerase (RPO), 90
 RNA polymerase occupancy, 117, 123, 145

- RNA-seq, 101–102, 104, 111–115, 117, 121–127, 129, 165, 183–184
- ROSETTA, 14
- RPO core enzyme, 90–91
- RPO holo enzyme, 90–91, 93, 95, 145
- run length encoding, 71
- Sanger sequencing/sequencer, 60, 67–68, 74–75, 80–81, 104
- scaffold, 7–8, 68, 74, 81–83, 109
- scripture, 107
- sense transcription, 125
- sequence comparison, 4, 51, 59
- sequence covariation, 12–13
- sequence similarity/identity, 14
- sequencing ancient DNA, 76
- sequencing depth, 61–62
- sequencing reads, 6, 66, 68, 72, 80–81, 85, 87, 102, 104, 107, 110, 195–196, 202
- Shigella*, 49, 51, 184
- shotgun sequencing, 1
- Sigma (σ) subunit or σ -factor, 90–91, 96, 145, 147, 150, 199
- signal transduction, 155, 170–171
- single input motif, 164
- single-molecule real time sequencing, 2, 200
- Smith–Waterman algorithm, 5, 70
- SMRT, 61–62, 64, 110, 200–204
- SNP, 45
- SOLiD, 60–62, 65–66, 83, 123
- solid phase PCR, 62
- Sorangium cellulosum*, 28
- Staphylococcus aureus*, 20, 30, 32, 46, 52–53, 78
- Staphylococcus* Cassette Chromosome, 53
- statistical model, 10
- stem loops, 12
- strand-specific sequencing, 105
- streamlining, 23–24, 29
- structural variations, 70, 76, 86
- sub-operonic transcriptional polarity, 126
- Sulcia mulleri*, 24–25
- supercoiling-sensitive genes, 136–137
- tandem repeats, 74, 81
- tetranucleotide frequency, 82–83
- tiling microarray, 126, 145, 151, 173, 175
- TMM (Trimmed Mean of M-values), 112
- topoisomerase I, 93, 134
- TransABYSS, 108
- transcript, 98–100, 105–107, 109–110, 122–124, 126–127, 129, 146
- transcript isoform, 100, 105, 109
- transcription, 5, 29–30, 85, 88, 90–93, 95–99, 117, 119, 121–122, 124–125, 127–134, 137, 144–150, 152–153, 155–156, 158–159, 164, 167, 174, 185, 187–188, 191, 193, 205
- transcription factor, 88, 117, 130, 158
- transcription initiation, 92–93, 95, 98, 117, 129, 131–132, 134, 145, 148, 188
- transcription regulatory network, 163
- transcriptional silencing, 180, 185
- transcription start site (TSS), 92, 124, 158, 205
- transcription termination, 98, 185
- transcription unit architecture, 122
- transcriptome, 99–102, 104–106, 108–122, 124–129, 133, 145–146, 148, 174–175, 178–179, 181, 183, 186, 190, 202
- transcriptome assembly, 116
- transfer RNA or tRNA, 12, 16, 90, 188
- transposon, 55, 75
- transposon insertion, 3, 152
- Trinity, 108
- two-channel microarray, 41
- two-component system, 169, 170
- unconstrained supercoiling, 93
- UP element, 92, 131
- UPEC, 18–19
- uropathogenic *E. coli* CFTO73, 18
- variant discovery, 45
- velvet, 69, 76, 81, 105, 108
- whole genome shotgun, 1, 6
- within-array normalisation, 41–43
- Yersinia pestis*, 75, 77
- Yersinia pseudotuberculosis*, 75
- Zero Mode Waveguides (ZMW), 200
- ZOOM, 70