

Index

- 1000 Fungal Genome Project, 221
 1000 genomes project, 310
 10k vertebrate genomes initiative, 49, 78
 16S rRNA. *See* ribosomal RNA
 18S rRNA. *See* ribosomal RNA

 28S rRNA. *See* ribosomal RNA

 454 GLS Flx. *See* Roche 454
 454 Life Sciences. *See* Roche 454

 Abyss, 312
 Acari, 36
 accessory genome, 113, 242
Acyrtosiphon, 36, 39, 42, 256
 adaptation, 169
Aegilops tauschii, 275
Agave, 275
 Agilent SureSelect, 142
Agitata, 197
 agriculturalism, 285
 alignment. *See* sequence alignment
 allelic variation, 285
 allopolyploidy, 132, 275, 321
 allotetraploid, 132, 275
 Alps, 131
 alternative splicing, 249
 Alzheimer's disease, 273
 Amaranthaceae, 272
Amaranthus hypochondriacus, 272
 Amazon, 169
 America, 105
 amino acid substitution matrix, 86
 amoeboid organisms, 195
 AmpliconNoise, 94
 ancestral states, 265
 ancient biomolecules, 282, 293, 298
 ancient divergences, 316
 ancient DNA, 7, 95, 179, 282–300
 analysis, 292
 authenticity of, 292
 laboratories, 291–2
 ancient metagenomic analysis, 297
 ancient RNA, 294, 297
Ancylostoma caninum, 75
 Animalia, 256
 annotation. *See* genome annotation
Anopheles aegypti, 255
Anopheles gambiae, 47, 255
 anthocyanin
 biosynthetic pathway, 270
 deposition, 268
 pigments, 269
 antibiotic resistance, 110, 242
 anti-malarial drug, 272
 Aphididae, 256
 Apiaceae, 277
 apomorphic genes, 237
 Applied Biosystems 5550 Genetic Analyzer, 43
 Apusozoa, 197
Arabidopsis thaliana, 230, 271
Arabidopsis, 276
 Araliaceae, 273
 Archaea, 245, 295
 archaeobacteria, 306
 archaeobotany, 282, 284, 286
 archaeoepigenetic analyses, 298
 archaeogenetic analysis, 286, 288
 archaeogenomics, 298
 Archaeognatha, 53
 Arctotideae, 266
 Área de Conservación Guanacaste, 176
 aRNA, 298. *See* ancient RNA
Artemisia annua, 272
 artemisinin, 272
 arthropod, 29

- Arthropoda, 252, 253, 256
 Articulata, 317
 Ascarididomorpha, 73
Ascaris suum, 72, 73, 75
 Ascidia, 95
 Asia, 105
 Asian carp, 187
 Asparagaceae, 275
 assembly. *See* genome assembly
 association mapping, 276
 association studies, 114
 Asteraceae, 264, 266, 272, 273, 275
 Australia, 105
 autolysis, 289
- Bacillus anthracis*, 242
 bacteria, 5, 101, 110, 163, 164, 243, 305, 306, 315
 hyperthermophilic, 245
 pan-genomes, 241
 bacteriophages, 104, 245
 bacterivores, 200
 baker's yeast, 224
 Baltic, 131
 banana, 274
 barcoding. *See* DNA barcoding
 barley, 275, 284, 285, 286, 297
 Barometer of Life, 183
 base composition. *See* DNA base composition
 Basic Local Alignment Search Tool. *See* BLAST
 Bay of Bengal, 109
 Bayesian, 71, 108, 166
 inference, 4
 methods, 111
 phylogenetic analysis, 109
 BEAST, 108
 bee-fly, 267
 Beijing Genomics Institute. *See* BGI
Bellis perennis, 126
 benthic macroinvertebrates, 182
 BGI, 48, 155, 167, 170
 BGI-10K, 49
 Bilateria, 253
 bioanalyzer, 293
 biodiversity, 14, 28, 62, 149, 154, 155, 160, 165,
 166, 168, 169, 171
 analysis, 175
 assessment of, 124, 177
 crisis, 188
 bioinformatics, 3, 6, 35, 89, 117, 154–65
 pipelines, 170, 171
 biological invasions, 187
 biological species concept, 196
 BioMarKs, 207
 biomolecular banks, 282, 288
 biomonitoring, 179, 186
- Biomonitoring 2.0, 180, 186
 bioprospecting, 186
 birds, 85
 birth-death process, 17, 44
 bitter melon, 273
 bivalves, 246
 BLAST, 16, 92–3, 139, 166, 238, 240, 242, 271
 expected values, 240
 homology, 271
 position-specific iterated, 240
 Blattaria, 30
 Blattodea, 30, 40, 53
Bombyx mori, 47
 bootstrap. *See* phylogenetic methods:
 bootstrapping
 Botanical Society of the British Isles, 138
Brachionus plicatilis, 88
 Branchiopoda, 29
 Brassicaceae, 273, 276
Brevimastigomonas, 199
 Britain, 131
 British Columbia, 107
Brugia malayi, 75, 243
Bursaphelenchus xylophilus, 75
- Caenorhabditis*, 317
 Caenorhabditis briggsae, 75, 243
 Caenorhabditis elegans, 62, 70, 72, 73, 74, 76,
 230, 232, 243
 development of, 76
 Cambridge, 106
Campylobacter, 242
 Canada, 107
 candidate gene, 147
 Cap3, 92, 93
 Carpathians, 131
 CAT model. *See* substitution models
 cave bear, 297
Cavernomonas, 198, 201
 CBD, 141
 CBOL, 142, 143, *See* Consortium for the Barcode
 of Life
 cDNA, 2, 38, 114, 200
 Cephalocarida, 29
 cercomonad, 196
Cercomonas, 197, 198
 Cercozoa, 196, 207
 character mapping, 266
 Chelicerata, 29, 43
 chemosensory, 45
Chlorophytum, 275
 Chromadoria, 65
 ciliates, 195
 CITES, 141
 cladogram, 161

338 INDEX

- Cleome*, 273
 climate change, 164, 170, 188
 clonal, 102
Clostridium difficile, 106
 clustering algorithms, 182
 Cnidaria, 246, 256
Cnidium dubium, 277
 COI. *See* cytochrome oxidase 1
 coalescence, 207, 223
 coalescent process, 320–3
 Codes of Nomenclature, 148
 codon bias, 108
 codon usage, 85
 Coelomata, 317
 COI. *See* cytochrome oxidase 1
 Coleoptera, 31, 36, 39, 40, 42, 49, 52
 collections
 archaeobotanical, 290
 herbarium, 282, 284, 286
 museum, 284
 Collembola, 29, 38
Colugo, 297
 comparative biology, 7
 comparative genomics. *See* genomics:
 comparative
 compensatory base changes, 204
 computational complexity, 20
 computational costs, 168
 concatenation, 316
 consensus. *See* phylogenetic method: consensus
 conservation, 50
 conservation genetics, 276
 conserved domains, 271
 Consortium for the Barcode of Life, 180, 210
 contamination, 290, 291, 309
 bacterial, 295, 309
 fungi, 292
 human, 291, 292
 microbial, 291
 plant, 292
 soil bacteria, 292
 yeast, 309
 contig, 91–3, 312
 convergent evolution, 109, 195
 Convolvulaceae, 273
Conyza canadensis, 273
 Copepoda, 29
 coprolites, 297
 core genome, 113
 Costa Rica, 176
 cotton, 295
 crop domestication, 285
 crop species, 276, 299
 crustacea, 29, 43, 86
 cryptic species, 134, 135
Cryptocercus punctulatus, 49
 Cucurbitaceae, 273
 cultivated crops, 285
 Cupressaceae, 277
 Curculionidae, 40
 Cuvier, Georges, 317
 cyanogenic glucosides, 274
 cyprinid, 88
 cytochrome oxidase 1, 4, 69, 70, 87, 88, 89, 95,
 142, 159, 179, 181, 205
 Czechoslovakia, 131

 D1–D2 variable regions. *See* ribosomal RNA
Dactylorhiza, 132, 133
 daisy, 264, 266
Daphnia, 35, 43, 235
 Darwin, Charles, 184, 223
 data analysis, 170
 data deluge, 220, 226
 data repositories, 167
 data sharing, 167
 data storage, 167–168
 dauer stage, 232
 dbEST, 37
 Denoising, 155
 Dermaptera, 31, 33, 53
 deuterostomes, 86
 developmental control genes, 230
 developmental genetics, 264
 developmental plasticity, 264
 diagenesis, 290
 diagnostics, 87
Dictyocaulus viviparus, 75
Dicyema, 88
 diploid, 132
 Diplura, 29
 Diptera, 31, 35, 36, 39, 40, 42, 44, 49, 52, 253
Dirofilaria immitis, 75, 77
 disabling mutations, 249
 discordance, 19
 disease management, 274
 diversification rates, 15
 DNA barcoding, 6, 7, 64, 66–70, 88, 123, 125,
 135–49, 159, 169, 176, 179–82, 188, 210,
 275, 297
 choice of locus, 67
 DNA base composition, 35, 73, 85, 164, 245,
 250, 308
 DNA fingerprinting, 125
 DNA fragmentation, 290, 293, 294
 DNA hybridization, 95
 DNA methylation, 145, 298
 DNA sequencing. *See* sequencing
 DNA-based identification, 136
 DNAnexus, 169

- domestication, 286
 genes, 284
 syndrome traits, 283, 284
 trait loci, 285
 traits, 284
- Doolittle, R. F., 237
- Dorylaimia, 65, 75
- Drosophila*, 35, 37, 44, 45, 46, 235, 239, 247, 251, 252, 254
Drosophila melanogaster, 230, 241, 243, 253
 segmentation, 230
- drug resistance, 109, 113
- early hominins, 300
- Ecdysozoa, 63, 317, 318
- Echinococcus*, 90
Echinococcus canadensis, 92, 93
- ecologists, 137
- ecology, 7, 129, 147, 149
 ecological adaptation, 324
 ecological observatories, 164
 ecological selection, 206
- Economic Botany Collection, 287
- ecosystems, 154, 162, 168
 functioning, 185
- ecotypes, 204
- Ectognatha, 29
- effective population size, 323
- eggplant, 276
- Embiodoptera, 30, 31, 42
- emergent properties, 185
- endogenous endonucleases, 289
- Endopterygota, 31, 253
- Enoplea, 78
- Enoplia, 65, 75
- Enoplida, 75
- enterobacteria, 245
- Entognatha, 29
- environmental microbiology, 202
- environmental selection, 285
- Eocercomonas*, 198, 201
- Ephemeroptera, 53
- epidemiology, 102, 106
- epigenetics, 145
- Escherichia coli*, 108, 114, 163, 232, 242, 246
- EST, 16, 37, 38, 45, 70, 243, 244, 247, 272, 275
 databases of, 71
- ESTs, 34, 37
- ESUs. *See* evolutionarily significant units
- eubacteria, 306
- Eukaryota, 253
- Eumetazoa, 253
- Euphorbiaceae, 273
- Europe, 105, 108, 129, 131, 132, 134
- Euryarchaeota, 255
- evo-devo, 7, 8, 229, 264, 265, 270, 271, 272
- evolution
 real-time, 284, 289
- evolutionarily significant units, 207, 208
- evolutionary developmental biology.
See evo-devo
- evolutionary ecology, 229
- evolutionary novelty, 229
- Evolutionary Placement Algorithm, 166
- exon, 37
- expressed sequence tag. *See* EST
- extinction risks, 183
- Fabaceae, 274, 276
- Felsenstein Zone, 22
- field taxonomy, 146
- fieldwork, 124, 128
 field biologists, 125, 139, 141, 143, 145
 field naturalist, 135, 138
 field-based sequencing, 142, 146, 147, 148, 149, 157
 field-based sequencing device. *See* sequencing: palmtop sequencer
- file formats, 311
 bam, 311
 bcf, 311
 Bed, 311
 cram, 311
 Fasta, 311
 Fastq, 311
 gff, 311
 gtf, 311
 sam, 311
 vcf, 311
- Filosa, 196
- FISH, 85, 88, *See* fluorescent *in situ* hybridization
- fish gut microbiota, 187
- fitness, 110
- flagellates, 195
- flavonoids, 268
- flax, 287
- florets, 267
- flow cytometry, 131, 145
- Flowsim, 91, 93
- Fluidigm access array, 142
- fluorescent *in situ* hybridization, 203
- food quality, 284
- food security, 283
- Formicidae, 39, 244
- forward genetics, 232
- fossil record, 282
- founder genes, 253
- fruit quality traits, 273
- functional annotation, 166
- functional genomics, 114

340 INDEX

- fungi, 7, 221
 metabolism evolution, 224
- GAL. *See* galactose pathway
- galactose pathway, 224
- gamma distribution, 318
- gastropods, 246
- GC content. *See* DNA base composition
- Genbank, 40, 124, 125, 139, 140, 141, 147, 205, 219, 222, 256
- gene
 cytochrome *b*, 283
 cytochrome P450, 272
LpoB, 246
Nv-osk, 241
Nwgl, 245
oskar, 239, 241
pbf, 285
pbf1, 286
Poldi, 249
su1, 285, 286
tb1, 285, 286
vsr1, 285
- gene clusters, 224
- gene discovery, 272
- gene duplication, 15, 17, 36, 42, 44, 314, 321
 tandem, 44
- gene evolution
de novo, 248
 duplication, 248
 horizontal gene transfer, 248
 reshuffling, 248
 retrotransposition, 248
- gene expression, 46, 47, 166
- gene expression homology, 271
- gene expression patterns, 271
- gene families, 17, 19, 36, 44, 314
- gene flow, 161, 323
- gene homology networks, 270
- gene loss, 17, 36, 44, 45, 321
- gene motifs, 271
- gene networks, 270
- gene order, 21, 85, 87
- gene silencing, 266
- gene tree, 18, 19, 33, 66, 315, 320
- gene tree parsimony, 19, 36
- generalized mixed Yule-coalescent, 207
- genetic code, 87
- genetic drift, 161, 250
- genome
 architecture, 224, 226
 assembly, 311
 filtering, 308
 reduction, 308
 reference, 293, 307, 308, 313
- Genome 10 K Project, 221
- genome annotation, 44, 50, 72, 95, 114, 163, 166, 167
 artefacts, 247
 RNA-seq and, 114
- genome assembly, 16, 48, 50, 72, 77, 92, 93, 103, 157, 167
 algorithms, 312
 assembly graph, 313
 De Bruijn graph assembler, 312
de novo, 114, 313
 draft, 160
 reference guided, 313
 scaffolding, 312
 software, 93, 312
- genome evolution, 42
- genome finishing, 160
- genome rearrangement, 73
- genome reduction, 42
- genome sequencing, 39, 42, 77, 78
 choice of targets, 48, 50, 53, 76
 large-scale initiatives, 54, 76, 77
 pre-publication availability, 76
- genome size, 39, 131
- genome-wide association studies, 234
- genomic selection, 276
- genomics
 comparative, 101, 103, 157, 305
 evolutionary, 305
 phylogenomics. *See* phylogenomics
- genotype–phenotype gap, 225
- genotyping, 308
- glissomonad, 196
- Globodera pallida*, 75
- Glycine max*, 17
- GMYC, 209, 210, *See* generalized mixed Yule-coalescent
- Google, 169
- Gorteria diffusa*, 264, 265–9
- Gorteriinae, 265–9
- Gossypium*, 296
- GPS, 124, 138, 140, 145
- Group B *Streptococcus*, 242
- Grylloblattodea, 30
- GTR. *See* substitution models
- GWAS. *See* genome-wide association studies
- Gymnadenia*, 131, 132, 145
- gyrA, 109
- Haemonchus contortus*, 75
- Haemophilus influenzae*, 101
- Haliotis asinina*, 247
- Hallwachs, Winnie, 176
- haploid, 73
- Harpegnathos*, 36

- Helicos Heliscope, 299
 Hemiptera, 36, 39, 40, 42, 44, 49, 52
 Hennig, 31
Heteromita globosa, 197
Heterorhabditis, 64, 75
 heterotrophs, 196
 heuristics, 20
 Hexapoda, 29, 38
 hexapods, 29
 hidden Markov model, 315
 high-throughput sequencing, 159
 HiSeq. *See* Illumina HiSeq
 historical samples, 290
 HMM. *See* hidden Markov model
 Holometabola, 31, 37, 40, 43
 Holozoa, 253
 homoplasy, 46, 73, 109, 110, 111
 homopolymer regions, 203, 206
 honey bee, 247
 horizontal gene transfer, 111, 224, 257
 horsetweed, 273
Hosta, 275
 Hox genes, 170
 Hubble, Edwin, 257
 human genome, 244
 Human Genome Project, 167
 human gut microbiome, 178
 Human Microbiome project, 167
Huperzia serrata, 272
 Huperziaceae, 272
 hybridization, 19, 23, 223
 hybridization capture, 297
Hydra, 235, 246
 hydrolysis, 291
 hydrolytic depurination, 290
 Hymenoptera, 31, 36, 37, 39, 40, 42, 44, 48, 49, 52

 i5k, 49, 78, 221
 Illumina, 43, 70, 72, 160
 GAI, 103
 HiSeq, 67, 155
 HiSeq 2000, 94, 95, 168
 HiSeq 2500, 43, 72, 160
 MiSeq, 43, 67, 91, 106
 NextSeq, 43
in situ hybridization, 265
 incomplete lineage sorting, 19, 22
 incongruence, 19, 22
 indels. *See* insertion–deletion events
 Indian Ocean, 232
 indicator taxa, 186
 industrial revolution, 287
 infectious diseases, 187
 Insecta. *See* insects

 insects, 5, 28–54, 253
 insertion sequences, 102
 insertion–deletion events, 276, 315
 insulin, 257
 interdisciplinary, 168
 internal transcribed spacer. *See* ribosomal RNA
 inter-retrotransposon amplified polymorphism, 287
 introgression, 19, 23, 223, 321
 intron, 37, 72
 invariant sites, 318
 invasive plant, 276
 Ion Torrent, 3
 PGM, 43, 106
 Proton, 106
 iPhy, 71
Ipomoea batatas, 273
 IRAP. *See* inter-retrotransposon amplified polymorphism
 Isoptera, 48
 ITS. *See* ribosomal RNA
 ITS1. *See* ribosomal RNA
 ITS2. *See* ribosomal RNA
Ixodes, 36, 43

 Janzen, Dan, 176
Jatropha curcas, 273
 Joint Genome Institute, 221
 Juncaceae, 276
Juncus effusus, 276
Juniperus tibetica, 277
 junk DNA, 296

 Kew Gardens, 287
 k-mer, 312

 La Réunion island, 234
 large subunit ribosomal RNA. *See* ribosomal RNA
 large-scale genomics initiatives, 49
 lateral gene transfer, 240, 257, 315, 318, 321
 Lepidoptera, 31, 36, 39, 40, 42, 49, 52
 LGT. *See* lateral gene transfer
 lice. *See* Phthiraptera
 likelihood, 4, 15, 20, 71, 105, 166
 limb regeneration, 247, 256
 lineage sorting, 223, 322
 lineage-through-time plot, 208
Linepithema, 36
 LINES, 219
 linkage maps, 276
 Linnaean classification, 299
 Linnaean species, 177
 Linnean binomial, 126–9
 LINUX, 72

342 INDEX

- Loa loa*, 75
Lolium perenne, 274
 long non-coding RNAs, 251
 long-branch attraction, 38, 86
 long-term ecological research, 185
Lotus corniculatus, 274
 LSU. *See* ribosomal RNA
 LTR-retrotransposons, 274
 LTT. *See* lineage-through-time plot
 lycophyte, 272
 lycopodium alkaloid biosynthesis, 272
 Lythraceae, 277
- maize, 244, 284, 285, 290
 Malvaceae, 274
 mammals, 85
 mammoth, 283, 296
 DNA, 297
 Mantophasmatodea, 30
 Map of Life, 183
 mapping, 103
 markers
 DNA barcode, 179
 microsatellite, 276
 phylogenetic, 306
 RAD-Tags, 219
 ribosomal, 179
 sequence tagged site, 274
 simple sequence repeat, 273
 SNP. *See* single nucleotide polymorphism
 maternal inheritance, 296
matK, 142, 143
 maximum likelihood. *See* likelihood
 maximum parsimony. *See* parsimony
 Mecoptera, 30, 31
Medicago sativa, 276
 Megaloptera, 30, 31
 MEGAN, 166
Megapalpus capensis, 267
 meiofauna, 62, 66, 69, 159
Meloidogyne, 75
 mesozoan, 88
Metabolomonas, 198
 metadata, 167
 metagenome, 163, 164, 166, 167
 metagenomics, 6, 157, 161, 165, 166, 219, 292,
 294, 295
 large-scale initiatives, 156
 metagenomic enrichment, 297
Metaseiulus, 36
 metasystematics, 183
 metatranscriptomics, 163, 165
 Metazoa, 253
Methanopyrus kandlerii, 245
 microarrays, 114, 160
 microbial eukaryotes, 207
 microfluidic, 164
 microRNA, 46–48, 219, 298
Microstegium vimineum, 276
 minicollagen, 256
 Mira, 92, 93
 miRNA. *See* microRNA
 missing data, 15, 16, 17, 18
 missing links, 299
 mitochondrial genomes, 33, 35, 70, 84–9, 93
 annotation of, 95
 genome structure, 88
 mitogenomics, 72, 84, 95
 mitogenomes. *See* mitochondrial genomes
 MLST. *See* multi-locus sequence typing
 mobile genetic elements, 104, 110, 113, 114
 mobile phones, 138
 molecular archaeobotany, 283
 molecular clock, 111, 283
 molecular dating, 108, 109
 molecular evolution, 102
 molecular operational taxonomic units, 67, 68,
 204
 molecular phylogenetics
 substitution models, 73
 molecular systematics, 125, 147
 history of, 4
 methods. *See* phylogenetic methods
 molecular typing, 102
 Mollusca, 86, 246
Momordica charantia, 273
 monophyletic, 134
 Moore's law, 222
 morphology, 129, 131, 134, 141, 147, 161, 164
 morphological characters, 30, 33, 62
 morphometrics, 132, 134
 morphospecies, 175
 most recent common ancestor. *See* MRCA
 moth, 274
 MOTU. *See* molecular operational taxonomic
 units
 MRCA, 108
 mRNA, 33, 35, 37, 70, 163, 169
 MRSA, 106, 109, *See* *Staphylococcus aureus*
 mtDNA, 84, 94, 95
 multi-locus sequence typing, 102
 multispecies coalescence methods, 322
Musa, 274
Musa acuminata cv. 'Calcutta 4', 274
 mutagenesis, 115
 mutation, 250
 mutation rate, 102, 107
Mycobacterium abscessus, 106
Mycobacterium tuberculosis, 102, 107, 110
Mycoplasma genitalium, 101

- Myriapoda, 29, 36
myxozoans, 256
- nacre formation, 247
Namibia, 266
NanoDrop 2000, 94
Nasonia, 36, 45, 241
National Center for Biotechnology information.
 See NCBI
natural historians, 137, 138
natural mutants, 269
natural selection, 109, 160, 161
 genome-wide scans, 324
NCBI, 76, 205, 271
ncRNA, 115
Neanderthal, 296
 genome, 297
Necator americanus, 75
nematodes, 5, 62–78, 175, 181, 182, 232, 243, 317
 protein space, 244
Nematomorpha, 63
Nematostella, 235
NEON project, 164, 165
Neotylenchinae, 64
neural network, 169
Neuroptera, 31
Newbler, 92, 93
Nigritella, 131
Nobel prize, 2
non-coding RNA, 115
non-genic regions, 251
non-synonymous sites, 108
novel gene evolution, 248
nuclear copies of mitochondrial genes. *See*
 NUMTs
nuclear genome, 74
nucleic acid integrity, 295
Nucleocercomonas, 199
nucleotide composition. *See* DNA base
 composition
nucleotide diversity, 90
NUMTs, 84
- Odonata, 40, 53
Olea, 273
Oleaceae, 273
oligonucleotide ‘baits’, 307
Oligostraca, 29
olive, 273
OMA, 314
Onchocercinae, 73, 74
operational taxonomic units, 162, 165, 178, 179,
 182, 204, 207, 210
Opisthokonta, 253
optical mapping, 157, 313
- orchids, 129, 130, 132, 138, 139, 140
orders of insects, 30
ORFans. *See* taxonomically restricted genes
orphan genes. *See* taxonomically restricted
 genes
orthologs, 35, 36, 70, 314
 algorithms for identifying, 314
 identifying, 314
 single copy, 314
OrthoMCL, 314
Orthoptera, 36, 40, 42
Oryza, 16
OTUs. *See* operational taxonomic units
Oxford Nanopore, 157
 GridION, 157
 MinION, 43, 144–6, 157
Oxyuridomorpha, 73
- Pacific Biosciences, 157
Paclitaxel, 272
palaeogenome, 296
palaeopolyploidy, 275
Paleoptera, 54
palmtop barcoding. *See* sequencing palmtop
 sequencer
Panagrellus redivivus, 75
Panax ginseng, 273
Pancrustacea, 253
pan-genome, 113, 242, 243
Paracercomonas, 198
parallel evolution, 109
paralogs, 44, 275
Paraneoptera, 40, 43, 54
parasites, 77, 195, 309
 eukaryotic, 306
 evolution of parasitism, 63, 64
parataxon, 148
parataxonomists, 137, 146, 148
Paraview, 170
parC, 116
parE, 116
Parhyale, 235
parsimony, 15, 19, 20, 105
 gene tree. *See* gene tree parsimony
pathogenicity, 101, 221
pathogens, 101, 102, 110, 117, 300, 309
 bacterial, 306
 plant, 286
 viral, 306
PCR. *See* polymerase chain reaction
pea, 274
pea aphid. *See* *Acyrtosiphon*
Pediculus, 39, 45
Pentastomida, 29
periculin genes, 246

344 INDEX

- permafrost, 295
petal spot, 268
 development, 264
Phasmatodea, 30, 31, 40
phenylpropanoids, 274
phenotype, 102, 147
Phlegmariurus carinatus, 272
Phthiraptera, 30
PhyloBayes, 74
phylodiversity. *See* phylogenetic diversity
phylogenetics, 108, 113
phylogenetic diversity, 40, 50, 51, 52, 53
phylogenetic incongruence, 316
phylogenetic methods, 4
 Bayesian inference, 316
 bootstrapping, 19, 20, 222, 223, 316, 317
 consensus, 20, 31, 321, 322
 maximum likelihood, 316
 parsimony, 316
phylogenetic substitution models. *See*
 substitution models
phylogenomics, 8, 14, 15, 23, 31, 35, 36, 71, 109,
 204, 219, 305, 314, 317
phylogeography, 108, 148
phylostratigraphy, 225, 252, 254
phylostratum, 253
phylotranscriptomics, 70
phylogenetic stage, 225
picobiliphytes, 164
Pinctada maxima, 247
Pisum sativum, 274
plantain, 274
plants, 8, 142
plasmids, 104
plastid, 164
Platanthera, 129, 130, 131
Platyhelminthes, 170
Plecoptera, 31, 53
plsC, 116
Poa, 275
Poaceae, 274, 275, 276
pocket sequencers. *See* sequencing: palmtop
 sequencer
Pogonmyrmex, 36
polar bear, 296
pollinators, 270
polyethylene glycol, 288
polymer consolidation, 288
polymerase chain reaction, 2, 4, 63, 67, 69, 86,
 115, 142, 159, 160, 169
 artefacts, 203
 chimaeras, 66, 69, 162, 182, 203, 205
 enrichment, 297
 error rate, 67
 inhibition, 179
 long, 35, 86, 88, 94
 multiplex, 297
 primers, 70, 86, 87, 88, 162, 182
 qPCR, 265
Polyneoptera, 33
polyphyletic, 134
polyploidy, 17, 42, 133, 275, 296
population genetics, 19, 134
population genomics, 88, 160, 161, 322
population size, 109, 161
ppplacer, 166
primates, 251
Pristionchus, 232
Pristionchus pacificus, 75, 232
professional taxonomists, 139, 146, 148
progenitor sequences, 248
prokaryotes. *See* bacteria
promoter regions, 114, 250
protein supercoiling, 245
protist, 7, 164, 204
 DNA barcoding, 210
 ecology, 195
 systematics, 195
 taxonomy, 195
protogenes, 250
Protostomia, 253
Protura, 29
Pscoptera, 30
pseudogenes, 224
pseudogenization, 44, 249
Psocodea, 30, 39, 52
Puccinia graminis subsp. *graminicola*, 274
pulse field gel electrophoresis, 104
QC. *See* quality control
QIIME, 161
QTL. *See* quantitative trait loci
quality control, 308
 filtering, 315
Quality Index for Predicted Proteins, 244
quantitative trait loci, 274, 275
quickly evolving species, 316
quinolone, 109
RAD. *See* sequencing
Raphidioptera, 30, 31
rapid evolution, 255
rare genomic changes, 219
rarefaction, 161
rat, 249
rate heterogeneity, 85
rbcl, 4, 142, 143, *See* rubisco
RdDM. *See* RNA-directed DNA methylation

- rDNA. *See* ribosomal RNA
- read mapping, 114
- recombination, 110, 111, 169, 182, 250
detection of, 111
hotspots, 111
- recorder effect, 177
- reduced-representation. *See* sequencing
- reference genome, 101, 104, 157, 164, 165–7
- regulatory genes, 266, 270
- Remipedia, 29
- repeat junctions, 275
- repeats. *See* repetitive DNA
- repetitive DNA, 36, 72, 103, 182, 275, 295, 296, 312
- repetitive elements. *See* repetitive DNA
- reproductive isolation, 148
- restriction sites, 308
- restriction-site Associated DNA sequencing. *See* sequencing
- reverse genetics, 232, 324
- Rhabdias bufonis*, 66
- Rhabditida, 75, 78
- Rhabditina, 78
- Rhabditomorpha, 73
- rhinoceros, 296
- ribosomal proteins, 38
- ribosomal RNA, 4, 5, 31, 85, 89, 162, 163, 169, 170, 257, 297, 306
- 16S rRNA, 163, 164, 179, 255
- 18S rRNA, 30, 51–3, 63–70, 163, 165, 196–200, 203, 205, 207
- 28S rDNA, 205
- D1-D2 variable regions, 205
- ITS, 16, 17, 18, 129–32, 142, 143, 145, 179, 180, 197, 199, 205, 297
- ITS1, 198, 205
- ITS2, 130, 204, 205
- secondary structure, 206
- V4 variable region, 205
- V9 variable region, 205
- rifampicin, 110
- RNA-directed DNA methylation, 298
- RNAi, 46
- RNA-polymerase, 110
- RNA-seq, 2, 8, 114, 275, 276, 308, 309
directional, 115
- Roche 454, 38, 46, 67, 68, 69, 70, 88, 89, 94, 103, 155, 160, 162, 202, 203, 205, 207, 209, 272, 273, 274, 275, 276
- Romanomermis culicivorax*, 75
- rotifer, 88
- rpoA*, 110
- rpoB*, 110
- rpoC*, 110
- rRNA. *See* ribosomal RNA
- rubisco, 179
- rye, 275
- ryegrass, 274
- Saccharomyces cerevisiae*, 224, 225, 250, 317
- Saccharomyces kudriavzevii*, 224
- salamanders, 247
- Salmonella enterica*, 108, 113, 114, 115
- Salmonella* Typhi, 115, 116
- Salvia sclarea*, 274
- sampling
artefacts, 202
chimaeras, 202
destructive, 290
environmental, 200
maximum information content, 177
voucher, 184
- Sanger, Frederick, 3, 257
- satellite model organisms, 230
- saturated, 35
- scaffolds, 157
- Scandinavia, 131
- scarab beetle, 232
- secondary metabolite pathways, 274
- secondary metabolites, 274
- secondary structure
RNA, 95
- self-organizing maps, 164
- sequence alignment, 4, 15, 16, 17, 18, 66, 71, 92, 93, 103, 111, 276
accuracy, 315
algorithms, 315
gold standard, 315
guide tree, 315
invariant sites, 318
multiple, 314
partitioning, 318
rate of evolution, 318
unreliable sites, 315
- sequence divergence, 16
- Sequence Read Archive, 40
- sequencing
capillary. *See* sequencing: Sanger
chimaera. *See* polymerase chain reaction:
chimaeras
cost of, 1, 72, 95, 144, 145, 155
depth of, 202
environmental, 6, 7, 160, 165, 167, 168, 199–204
errors, 64, 67, 69, 91, 144, 155, 158, 162
EST, 305
Illumina, 160, 203, 273
large-scale initiatives, 156
magnetic, 159
multiplex, 88

346 INDEX

- sequencing (cont.)
 multiplexing, 103
 nanopore, 144
 palmtop sequencer, 123, 124, 137, 139, 143, 144
 platforms. *See also* Illumina, Roche, Pacific Biosciences, Ion Torrent, Oxford Nanopore, Helicos, Applied Biosystems
 platforms, comparison of, 158
 pooled samples, 95
 RAD, 267, 274, 275, 276
 reduced-representation, 305, 307, 313
 Sanger, 3, 4, 67, 88, 95, 125, 155, 159, 202, 219, 297
 short-read, 307
 shotgun, 292, 294, 295
 single-cell, 78, 163, 164
 single-molecule, 3, 144, 157, 159, 299
 tags, 308
 targeted, 5, 142, 160
 taxonomic assignment, 166
 technology, 6, 125, 155
 transcriptome. *See* transcriptome
 whole-genome, 71, 147, 307
- shell evolution, 247
- Shigella*, 109, 163
- short interfering RNA, 298
- signature proteins, 255
- simulation, 91–93, 170
- SINEs, 219
- single nucleotide polymorphism, 103, 108, 110, 111, 168, 275, 276, 291, 293, 296, 308, 309
- single-molecule sequencing. *See* sequencing: single-molecule
- Siphonaptera, 30, 31
- siRNA. *See* short interfering RNA
- size of phylogenetic trees, 13, 14
- small RNAs, 43, 46
- small subunit ribosomal RNA. *See* ribosomal RNA:18s rRNA
- SNP. *See* single nucleotide polymorphism
- SOAPdenovo, 312
- Solanaceae, 276
- Solanum melongena*, 276
- SOLiD, 43
- Sonneratia*, 277
- South Africa, 266
- South America, 105
- Southern Tibetan Plateau, 277
- species circumscription, 126–35, 140, 141, 147–9
 molecular, 143
- species description, 126
- species diversity, 40, 161
- species identification, 124
- species richness, 182, 186
- species tree, 19, 33, 66, 84
- Sphaerulariidae, 64
- Spiruria, 64, 70, 74
- Spirurina, 73
- Spiruromorpha, 73
- split network, 31
- SplitsTree, 32
- Sporozoa, 195
- SSU. *See* ribosomal RNA
- Staphylococcus aureus*, 104, 108, 114, 242
- staple crops, 286
- Steinernema*, 64
- Strepsiptera, 31, 39, 53
- Streptococcus agalactiae*, 242
- Streptococcus pneumoniae*, 102, 111, 112, 113
- stress-resistant plants, 272
- Strigamia*, 36
- Strongyloidoidomorpha, 64
- Strongylomorpha, 64, 70
- substitution models, 4, 86, 319
 Bayesian MCMC, 320
 CAT model, 319
 covarion, 320
 general Markov, 320
 General Time Reversible, 318
 Jukes–Cantor, 318
 mixture models, 320
- substitution rate, 107
- supermatrix, 15, 16, 19, 20, 316, 319, 321
- super-spreader, 107
- supertree, 16, 19, 321
- surrogate taxa, 186
- sweet potato, 273
- Switzerland, 277
- symplesiomorphy, 252
- synapomorphy, 46, 87, 252
 synapomorphic genes, 239
- synonymous sites, 108
- systematic error, 320
- systems biology, 186
- T4 bacteriophages, 245
- Taenia*, 92, 93
- tapeworms, 89
- Tara Oceans, 167
- Tardigrada, 30
- target enrichment, 296, 297
- Tasmanian tiger, 296
- Taxaceae, 272
- Taxol, 272
- taxon delimitation, 196

- taxonomically restricted genes, 8, 17, 257
 classification, 239
 evolution, 248, 250
 functional significance, 244
 origins, 248
 sampling error, 240
 similarity threshold, 240
 systematics of, 252
 ubiquity of, 241
- taxonomically unique sequences. *See*
 taxonomically restricted genes
- taxonomy, 6, 124–8, 135–7, 147, 148
 Codes of Nomenclature, 126
 taxonomic impediment, 137
 taxonomic lumping, 199
 taxonomic splitting, 199
- Taxus mairei*, 272
- TB. *See* tuberculosis
- Tenebrionidae, 40
- terpenoids, 274
- Tetranychus*, 36
- Thailand, 104, 106
- Theobroma*, 274
- ThermoFidelase sequencing kit, 246
- third generation sequencing, 299
- Thysanoptera, 40
- ToL. *See* tree of life
- tooth fragments, 295
- Topo V, 245
- topoisomerase, 245
- total evidence, 321
- TraDIS. *See* transposon directed insertion-site
 sequencing
- Tragopogon miscellus*, 275
- transcriptome, 16, 37, 43, 49, 70, 71, 74, 114
- transcriptomics, 8, 45, 70, 166, 270, 272, 273
 comparative, 271
- transmission, 108, 110
- transposable elements, 37
- transposon directed insertion-site sequencing,
 115
- transposons, 104
- tree of life, 23, 219–26, 318
- tree-thinking, 221
- Tremula*, 197
- Tribolium castaneum*, 42, 235
- Trichinella spiralis*, 75
- Trichoptera, 31, 40
- Trichuris*, 75
- tRNA, 85, 95
- trnL-F*, 143
- trophic structure, 185
- tuberculosis, 107
- turfgrass, 274
- Tylenchomorpha, 70
- Typha minima*, 277
- Typhaceae, 277
- ultra-barcoding, 274
- uncultured, 166, 169
- universal barcoding gene, 210
- universal primers, 203
- unstable taxa, 316
- untranslated regions. *See* UTRs
- UTRs, 115
- V4 variable region. *See* ribosomal RNA
- V9 variable region. *See* ribosomal RNA
- vaccine, 111
- variable number tandem repeats, 102, 104
- Vavilov Institute, 287
- vectorial capacity, 221
- Velvet, 312
- Vibrio cholerae*, 109
- viral genomes, 243, 245
- viral transfer, 243
- virosphere, 243
- virus reservoirs, 243
- viruses, 164
- VisTrails, 170
- VNTRs. *See* variable number tandem repeats
- voucher specimen, 95, 139–41
- wasp, 241
- waterlogged timber, 288
- wheat, 275, 284, 286, 290
- whole genome duplication, 17, 42
- whole genome sequencing. *See* sequencing:
 whole-genome
- wiki, 76
- Wilson, Edward O., 175
- Woese, Carl, 255
- WormBase, 77
- yeast, 239, 317
- Yersinia pestis*, 295
- Zootermopsis nevadensis*, 39
- Zoraptera, 31, 33, 53
- Zygaena filipendulae*, 274