Index

Acetolactate synthase, 263
Acetyl-CoA, 168, 252, 285
Activation, cassette-type, 172
Adaptation, 10–11
Affymetrix GeneChips, 142
Aldolase, 262
Algorithms
 alignment, 229
 elementary mode analysis, 256–257
 MACAW, 136
 for motif discovery, 135–140
 for protein categories, 239
 sequence alignment, 265
AlignACE, 129, 136–142
Alignment
 algorithms, 229
 of metabolic pathways, 259–263, 265
Amino acid production, 59
Amphiphilic molecules, 268
Ancestry, and homology, 6, 7
Anemia, hemolytic, 279–280
Annotation
 five step approach, 24
 of new genome (*E. coli*), 103–104
 of pathways, in genomes, 263–267
 pitfalls, 20–21
 specific systems, 21–24
 vocabulary, 33
Antigens, cell surface, 163
Apoptosis
 and Boolean network, 197, 201–202
 and cancer cells, 212
 as cell state, 183
 and Ras, 189
Archaea
 glycolysis, 263
 proteomes, 26, 34
Arginine biosynthesis, 133
Arrays. *See* Microarrays
Attraction, basins of
 and cancer, 211
 and cell state change, 201, 203, 204–205
 description, 197, 199–200
Attractor states, 199–202, 203–205
 and cancer, 210–212
 pseudo, 208–209

B. burgdorferi, 266–267
B. subtilis, 121, 124
 purine biosynthesis, 132
Bacteria. *See also specific bacteria*
 chemotaxis, 280–281
 DNA-binding regulators, 114
 DNA regulatory motifs, 139–142
 glycolysis, 262–263
 operon prediction, 121–124
 proteomes, 26–27, 34–35
 transcriptional units, 116–117
Bacteriophage lamda, 282–283
Basic Local Alignment Search Tool (BLAST), 223, 229
Bayes theorem, 89
Binding sites
 predicting, 87
 for transcription factors, 94, 98, 129
Biochemical analysis, 260–263

BioKleisli, 57–58
Biological noise, 208–209
Biomechanics, 6
Biopathways Consortium, 213
Biosynthesis
 cholesterol, 167–169
 heme, 162–163
 modeling, 64
 nucleotide, 263
 purine, 132
Birds, wings, 4
BLAST. *See* Basic Local Alignment Search Tool
Boltzmann equation, 88–89
Boolean networks
 architecture, 194–196, 205–207
 attractor states, 199–202, 203–205, 208–209, 210–212
 basins of attraction, 197, 199–200, 201, 203, 204–205
 and cancer, 181, 189, 209–212
 and cell states, 196–198, 200–209
 directionality, 199
 dynamics, 196–198
 and expression profiles, 213
 limitations, 212–213
 on-off switching, 192–194
 pleiotropy, 205–207
 protein-gene interaction, 213
 proteolysis, 195–196
 state space, 197, 198–200, 203, 214
 trajectories, 199–200, 202
 unstable states, 198, 199, 203
 updating, 199
BRENDA, 49, 57, 66, 252
Burns, 267
Butterflies, eyespots on, 3

C. acetobutylicum, 132
C. glutamicum, 263
Cadherins, 231
Caenorhabditis elegans
 comparative analysis, 36–41
 repeating domains, 40
Cancer, and Boolean networks, 181, 189, 209–212
Carbon dioxide transport, 279
Cassette-type activation, 172
Catalysts, 268
Celera Genomics, 76
Cell membrane
 and cholesterol, 160
 in simulated cell, 274, 278, 279
Cell(s)
 construction of, 59
 virtual, 277–280 (*see also* E-CELL)
Cell Signaling Networks Database, 213
Cell states. *See also* Differentiation; Signal transduction pathways
 in Boolean network, 196–198, 200–209
 and cancer, 209–212
 and context, 208–209
 definitions, 182–183
 and gene network hierarchy, 172–173
 perturbations, 189–190, 201–209
 robustness, 202
 in rule-based model, 60
Cell surface antigens, 163
Chance, 6, 7, 8
Chemical kinetic simulation, 164–167
Chemotaxis, 280–282
Cholesterol, 155–156, 158–160
 mathematical model, 167–169
Cleavage sites, predicting, 27
Clustering
 as analysis approach, 27–32
 coinheritance groups, 231
 databases, 27–28
 DNA regulatory motifs, 140
 and expression profiling, 213
 hierarchical, 29
 on mRNA expression, 131
 by regulatory region, 234
 of related enzymes, 260
 and transcription regulation, 116
Clusters of Orthologous Groups (COGs), 27–30, 260, 265
CluSTr, 27–32, 34
Co-A transferase, 133
Coinheritance groups, 228–231
Communication processes, 64–65. *See also* Signal transduction pathways
Comparative analysis
 and elementary modes, 257, 263

on eukaryotic proteomes, 34–42
of genomes, 103–104, 232, 233
and RNAP binding, 110–117
Concentration rates, in models, 63, 65
velocities, 166–167
Congruence, 117–120
Conservation
of DNA regulatory motifs, 141
of genes in operons, 122–124
and phylogenetic space, 240
Context
and cell states, 208–209
and protein function, 223–224, 225, 226–234, 239–240
Contingent events, 7
Convergence
and homology, 6, 7, 8
in signal transduction, 187–189
Cooperativity, 193
CORBA (Common Object Request Broker Architecture), 74, 76

Database Management Systems (DBMS), 51–52
Database of Interacting Proteins (DIP), 238
Databases
biochemical, 49, 252, 261
classification, 29–32
for clustering, 27–28
vs. DBMS, 50–52
enzymes, 54
EST, 235
for functional groups, 134–135
gene network, 150
gene regulation, 54
genomic neighbors, 231
integrating, 54–58, 69–76
listing, 54
microarray, 235
pathways, 49, 54, 213–214, 238
promoters, 234
protein interactions, 238
quality control, 55
redundancy, 29–30
remote access, 73
Saccharomyces, 33
search algorithms, 28
sequence annotation, 20–25
sequences, 27–28, 49, 53, 54, 228, 264
transcription factors, 49
transcriptome experiments, 105
user queries, 30–32
Data merging, 55
Data Warehouse, 57
DBMS. *See* Database Management Systems
Differentiation
and Boolean network, 197, 200–202
as cell state, 182–183
gene network control, 160–163
induced, in culture, 190
vs. quiescence, 183
Digitalization. *See* Boolean Networks
Dihydroxyacetone kinase, 265
DIP. *See* Database of Interacting Proteins
Distal-less gene, 4, 10
Distance, intergenic
and coinheritance, 229
and functional relatedness, 116, 121, 122, 124, 132, 229, 233–234
in regulatory motif search, 139–140
in *S. cerevisiae*, 142
from start codon, 139–140
Distribution, hypergeometric, 139
DL-diaminopimelate, 263
DNA-binding proteins, 87–93, 98. *See also* Weight matrices
DNA microarrays, 235
DNA regulatory motifs, 139
DNA replication factors, 241
DNA sequences
motif discovery within, 135–140
similarity, and homology, 5
Domain fusions, 225–228
Domains
annotation, 25, 34
C2H2-type zinc finger, 37–38, 40, 41
classification, 29
collagen triple helix repeat, 39
comparative distribution, 38
composition, 41
C-type lectin, 39
C4-type steroid receptor zinc finger, 39
EGF-like, 40, 41
fibronectin type III, 40

Domains (cont.)
immunoglobulin, 40
LDL-receptor, 40–41
ligand-binding, of nuclear hormone receptor, 39
major histocompatibility complex, 40
nonspecific, 34
repeated across sequence, 40
Rosetta Stone links, 225–228
Drosophila melanogaster
comparative analysis, 36–41
domain repetition, 40–41
Dyad-detector method, 112

E-CELL
cell models, 277–280
gene regulation model, 282–283
mitochondria model, 283–285
scope, 273–275
signal transduction model, 280–282
software, 59
EcoCyc, 238
Electroneutrality, 279
Elementary modes, 255–259, 263
EMBL. *See* European Molecular Biology Laboratory
EMPATH, 257
Encephalopathy, mitochondrial, 285
Energy. *See also* Glycolysis
binding, 87–89, 90–93
and Boolean networks, 198–199
and enzyme types, 252
and mitochondria, 285
multicellular organisms, 181–182, 198–199
Engrailed expression, 3, 10
Entner-Doudoroff pathway, 262
Entrez, 22–23, 231
Entropy, relative, 89, 93
Environmental modulation, 208–209
ENZYME database, 54, 252
Enzymes. *See also specific enzymes*
and cholesterol, 159–160, 168
databases, 54
defined, 60
elementary modes, 255–259, 263
heme biosynthesis, 162–163
hierarchies, 263
Michaelis-Menten kinetics, 193
modeling, 59, 66–67
nitrite reductase, 227
pH dependence, 279
types, 252–253
EPD. *See* Eukaryotic Promoter Database
Erythrocytes
as irreversible process, 160–163
models, 169–171, 278–280
and pyruvate kinase deficiency, 279–280
Erythroid progenitor cells, 154
Erythropoietin, 160, 170–171
Escherichia coli
acetate Co-A transferase, 133
acetate synthesis, 263
adaptation, and congruence, 118–119
chemotactic response, 280
coregulated genes, 112–113, 119–120
DNA regulatory motifs, 141
and domain fusion, 227
flux conditions, 259
gene conservation, 123
gene regulation simulation, 282
glycolysis, 262, 282
lactose, 182
metabolic model, 78
phosphorylation, 281
promoters, 107–116
purine biosynthesis, 132
transcription, 104, 113–114, 116–117
EST. *See* Expressed Sequence Tags
Eukaryotes
and operons, 133
phylogenetic diversity, 228
Eukaryotic Promoter Database (EPD), 54, 234
Eukaryotic proteomes
cluster data, 29
comparative analysis, 35–41
complete sets, 26–27
domain distribution, 38
phylogenetic classification, 27–28
RNA binding, 38
European Molecular Biology Laboratory (EMBL), 49, 53, 54
E-values, 265

Events, reaction *vs.* regulatory, 152–153, 155–156
Evolution. *See also* Phylogenetic profiles
 and Boolean network, 205–207, 208
 and transcription regulation, 105, 121, 124
ExPASy. *See* Expert Protein Analysis System
Expectation-maximization (EM) method, 96
Expected values (E-values), 256
Expert Protein Analysis System (ExPASy), 54, 261
Expert systems, 59
Expressed Sequence Tags (EST), 235
Expression
 basic elements, 282
 of distal-less gene, 4, 10
 E-CELL model, 282–283
 mRNA, 130–131, 142–143, 235–236
 profiling, 213
 and promoter specificity, 87
 and protein function, 234–237
 from proteomes, 236–237
 RegulonDB search, 117
Eye, homology, 8–9
Eyespots, of butterflies, 3

False positives, 106, 109, 111–112
Feedback
 and Boolean network, 193
 cholesterol, 158–160, 168–169
 erythrocytes, 161, 162, 171
 positive and negative, 172
Flight, 11
Flowers, petals of, 5
FlyBase, 33
Footprinting, 120, 135, 140
Fossils, Burgess shale, 7
Functional dynamics, 163–167
Function(s). *See also* Protein functions
 and DNA motifs, 129
 of network types, 156–163
 relatedness, 133–135
 and transcription regulation, 114
Fungal protein, 38
Fusion, domain, 225–228

Galactose utilization, 142
GATA1, 170–171
GATA1-p, 154, 158, 161, 162
GDB. *See* Genome Database
GenBank, 54, 228
GeneNet, 49, 150–158, 168
Gene Ontology (GO), 33, 34
Genes. *See also* Expression
 coexpression, 112
 context, 225, 226–234, 239–240
 coregulated, 112–113, 119–120, 140–141 (*see also* Regulons)
 distal-less, 4, 10
 distance between, 116, 121, 122, 124, 132, 229, 233–234
 functionally related, 133–135
 homeobox, 10
 homology, 8
 within operons, 117–118, 121, 122–123
 regulatory, 8
 relative positions, 231–233
 signaling, 184
 transition-initiating, 207
Genomes. *See also* Proteomes
 comparative analysis, 35
 human, 35
 pathway annotation, 263–267
 and protein function, 225–234
 protein networks in, 240–243
 relative positions in, 231–233
 signal transduction portion, 184
 starting new, 103–104
Genome Database (GDB), 54
Gibbs′ sampling, 96–97, 112, 136
 in AlignACE, 136–137
Gluconeogenesis, 267
Glucose, lactose, RNA-polymerase, 63–64
Glucose-6-phosphate, 60–61
Glycolysis
 E. coli simulation, 282
 elementary modes, 258–259
 pathway alignment, 262–263
 rule network, 67–68
 in simulated cells, 278–279
G-protein coupled receptor, 38
Greedy algorithms, 95

Growth and development, 160–163
Growth factors, 183–184, 189

Haemophilus influenzae, 116, 259
Heat shock response, 142
Hedgehog protein, 3
Helicobacter pylori, 262
Helix-turn-helix, 114, 121
Hemoglobin, 193, 279
Hemolysis, 279
Hemolytic anemia, 279–280
Hidden Markov Models (HMMs), 23
Homeobox genes, 10
Homeostasis, 158–160, 167–169
Homology
 benefits, 11–12
 definitions, 2–3, 5–6
 at differing levels, 3–4, 9–10
 and dissimilarity, 5
 distal-less gene, 4, 10
 DNA sequences, 5
 gene expression patterns, 4
 of gene vs. end product, 8–10
 as hypothesis, 3, 6–8
 limitations, 223–224
 vs. paralogy, 6
 of regulatory circuits, 1–2
 serial, 5–6
 and surrounding parts, 6
Homology-derived secondary structure of proteins (HSSP), 27, 31
HTML. *See* HyperText Markup Language
Human genome, 35
Human Genome Project, 53
Human Genome Sciences, 76
HUSAR, 57
Hypergeometric distribution, 139
HyperText Markup Language (HTML), 53, 73

Immunoglobulin domain, 40
INCYTE Genomics, 77
Information fusion, 56, 74
Information system(s)
 defined, 52
 signal transduction as, 181–182, 186–188, 205–206
Informax, 77
Insect cuticle protein, 38
Internal congruence, 119
InterPro
 and CluSTr, 34
 for comparative analysis, 35, 37–41
 controlled vocabulary, 33–34
 description, 23
 and domain composition, 41
 and GO, 34
 protein analysis, 27
 user interface, 33

Jaccard coefficient, 239
JAVA, 73–75, 156
JDBC (Java DataBase Connectivity), 58, 73–74

KEGG. *See* Kyoto Encyclopedia of Genes and Genomes
Key word recovery, 238–239
Kullbach-Liebler distance, 89, 93
Kyoto Encyclopedia of Genes and Genomes (KEGG)
 and database integration, 57
 and DNA regulatory motifs, 141
 and functional relationships, 238
 interface, 53
 and metabolic pathways, 49, 134–135, 252, 261
 url, 54

Lactate dehydrogenase, 253
Lactic acidosis, 285
Lactose, 63–64, 182
LASSAP, 29
Leigh syndrome, 285
LexA, 111–112, 113
LIGAND database, 54
Limbs, 4, 5
Lion Bioscience, 77
Lipidation, 237
Lipids
 biosynthesis, 266–267
 metabolism, 155–156
Liver, 267
Low-density lipoproteins (LDLs)

domain, 40–41
role, 160, 167–168
Lysine, 263
Lytic-lysogenic switch, 282–283

MACAW. *See* Multiple Alignment Construction and Analysis Workbench
Major histocompatibility domain, 40
Mammals, 172–173
MAP (mitogen-activated protein) kinase, 185, 187, 193
cascades, 208
MAP score, 138
MARGBench. *See* Modeling and Animation of Regulatory Gene Networks
Markov models, hidden (HMMs), 23
Mating type, 142
Matrices. *See* Elementary modes; Weight matrices
MEME. *See* Multiple EM for Motif Elicitation
Metabolic flux space, 259
Metabolic networks
diagrams, 59
glucose-6-phosphate, 60–61
glycolysis pathway, 67–69
and homology, 223
MARGBench example, 56, 69–76
mathematical models, 58–59
Metab-Sim simulation, 58, 66–69, 71
in multicellular organisms, 182
rule-based modeling, 60–66, 76–79
simulators, 163–167
Metabolic pathways
alignment, 259–263, 265
boundaries, uncertain, 267
branched and interconnected, 252–253
catabolic and anabolic, 252
defined, 252–253
genome analysis, 263–267
identification
elementary modes, 255–259, 263
by synthesis, 255
and organism complexity, 268
Metabolism
comparative analysis, 41–42
proteome proportion, 41–42
Metabolites
identification, 257–258
in models, 60, 62–63, 65
MetabSim, 58, 66–69, 71
and MARGBench, 75–76
Methanococcus jannaschii, 263
Metrics, of protein function, 238
Mevalonate pathway, 159–160
MGD. *See* Mouse Genome Database
Michaelis-Menten kinetics, 66–67, 165, 167, 193
Microarrays, 235
Millipedes, 5
Minimal upstream regions, 116
MIPS. *See* Munich Information Centre for Protein Sequences
Mitochondria, 283–285
Modeling and Animation of Regulatory Gene Networks, 56, 69–75
application, 75–76
Models. *See also* Boolean networks; Simulation
chemical-kinetic, 164–167
of cholesterol biosynthesis, 167–169
gene regulation, 65
mathematical, 150, 191–192
of metabolic pathways, 267–268 (*see also* Metabolic pathways)
rule-based, 60–66
Modularity, 190–191
Molecular databases, 53–58
Morphogenesis, 160–163
Motifs
common, 112
discovery process, 129, 130
DNA-binding, 114
DNA regulatory, 135–140
overlapping, 112–113
palindromic, 140
RNA-binding, 38
Mouse Genome Database (MGD), 33
mRNA expression, 130–131, 142–143, 235–236
Multiagents, 57
Multicellular organisms, 181–182, 209–212

Multiple Alignment Construction and Analysis Workbench (MACAW), 136
Multiple EM for Motif Elicitation (MEME), 136
Munich Information Centre for Protein Sequences (MIPS), 134–135, 142, 238
MURS. *See* Minimal Upstream Regions
Muscle development, 9
Mutations
 and Boolean networks, 210–212
 in simulated cell, 281
Myc, 189, 208
Mycobacterium tuberculosis, 244–245
Mycoplasma genitalium, 265, 277–280, 282
Mycoplasma hominis, 262
Mycoplasma pneumoniae, 264–265, 266
Myopathy, mitochondrial, 285

Nematodes, 38
Neostructuralists, 6
Network(s). *See also* Metabolic networks
 cholesterol biosynthesis, 167–169
 components, 172
 congruence, 119
 control, 152
 defined, 49, 149
 diversity, 163–164
 elementary structures, 151
 functional dynamics, 163–167
 functional types, 156–158
 gene regulation, 282
 homeostasis-maintaining, 158–160
 and irreversible processes, 160–163, 169–171
 lytic-lysogenic, 282–283
 mammalian, 172–173
 in multicellular organisms, 182
 object oriented analysis, 151–158
 organization, 150–151, 163, 172–173
 protein interaction, 240–243
 signal transduction, 188, 190–191 (*see also* Boolean networks)
 transcription initiation, 104, 106–117, 124
Nitrite reductase, 227
Noise, biological, 208–209
Nucleotides, 263, 278

Object-oriented systems, 59, 66
 for gene networks, 151, 163, 164–167, 172
Object Query Language (OQL), 74
ODBC. *See* Open Database Connectivity
Oligonucleotides, 136
Open Database Connectivity (ODBC), 73
Open Reading Frame (ORF), 103–104, 139
Operator sites, 106–107
Operons
 in bacteria, 121–124, 132, 141
 in *E. coli*, 104, 116–117
 in eukaryotes, 133, 231
 and evolution, 122–124
 genes within, 117–118, 121, 122–124
 lac and *ara*, in *E. coli*, 282
 and protein function, 231–233
 and regulatory motifs, 141
OPM, 57
OQL. *See* Object Query Language
ORF, ORFs. *See* Open Reading Frame
Organisms, multicellular, 181–182, 209–212
Organs, morphogenesis, 160–163
Orthologues, 38, 116. *See also* Clusters of Orthologous Groups (COGs)
Osmosis, 279
Oxidation, 237
Oxygen transport, 279

Palindromes, 140
Paralogy, 6
Pathways. *See also* Signal transduction pathways
 databases, 54
 liver antioxidant, 267
 MAP kinase, 185, 187, 193
 metabolic (*see* Metabolic pathways)
 modeling, 65
 novel, 244–245, 253
 ribulose, 266
PDB. *See* Protein Data Bank
Pearson correlation, 140
PEDANT, 57
Pentose phosphate cycle, 258–259, 262, 278, 282
Peptides, signal, 27

Petri nets, 59
pH, 279
Phosphofructokinase, 59, 262
Phospholipid biosynthesis, 274, 278
Phosphorylation
 and Boolean networks, 195–196, 198–199
 and external stimuli, 280–282
 and metabolic pathway, 252
 and protein modification, 237
Phylogenetic footprints, 120, 135, 140
Phylogenetic profiles, 115, 134, 228–231
Phylogenetic space, 240
PIR. *See* Protein Information Resource
P13K-Akt sequence, 185
Pleiotropy, 187–189, 205–207
Prediction
 of cell regulative disorders, 214
 of protein-binding sites, 106
 of protein functions, 237–239
 of regulatory network, 104, 106–117
 of regulons, 130–135
 of signal peptides, 27
Predictive value, 109
Profiling, gene expression, 213
Proliferation, 181
Proliferation, cellular, 197, 211–212
Promoters
 and discretization, 194
 in *E. coli*, 103, 107–117
 predicting, 106–107, 109
 and RNAP, 106–117
 specificity, 87, 94, 98
PROSITE database, 22
Protease, sterol-regulated, 155, 160
Protein Data Bank (PDB), 31, 54
Protein functions
 categories, 34, 41, 239–240
 and expression, 234–237
 and gene position, 231–233
 measuring, 237–239
 metrics, 238
 molecular *vs.* cellular, 224
 phylogenetic profiles, 228–231
 Rosetta Stone approach, 225–227
 structure-function relationship, 266
Protein fusion, 225–228
Protein-gene interactions, 114, 213
Protein Information Resource (PIR), 49, 54
Protein interaction networks, 241–243
Protein kinases, 37–38
Protein-protein interactions
 in Boolean networks, 195–198, 199
 and homology, 223–224
Proteins. *See also* Yeast
 annotation, 20–25
 chicken, 154
 classification, 26–34, 41
 constraint sharing, 225
 DNA-binding, 87–89
 eukaryotic (*see* Eukaryotic proteomes)
 fungal transcription, 38
 fusion, 225–228
 Hedgehog, 3
 human, 155
 insect cuticle, 38
 lengths, 36–37
 master, 208
 modifications, 237
 multidomain, 41
 Myc, 189, 208
 Ras, 189, 208, 241
 reading frames, 264
 regulatory, 106, 189
 rhodopsin-like, 38
 RNA-binding, 38
 secreted, 27, 37
 species-unique, 38
 SREBP, 155, 159–160, 168
 synthesis simulation, 278
 transition-causing, 208
 transmembrane, 27, 37
Proteolysis, 195–196
Proteomes. *See also* Comparative analysis; Eukaryotic proteomes
 analysis, 25–33
 annotation, 20–25
 of bacteria, 26–27, 34–35
 classification, 25–33
 comparative analysis, 34–41
 expression profiles, 236–237
 functional analysis, 34, 41
 sequencing history, 19–20

ProtoMap database, 28
Pseudo attractors, 208–209
PSI-BLAST, 265
Purine biosynthesis, 132, 133
Purine metabolism, 59
Pyruvate kinase, 262, 279–280

Quality, of data, 55
Queries, 30–33, 55, 57–58. *See also* Searches

Ras, 189, 208, 241
Reaction events, 152, 155
Reading frames
  protein- vs. RNA-coding, 264
  reannotating, 265
Reannotation, 264–267
Receptors, 7-helix G-protein coupled, 38
Redox balance, 257–258
Redundancy
  in databases, 29–30, 51–52
  in signaling, 187
Regulation, in multicell organism, 182. *See also* Feedback; Transcription factors
Regulatory circuits, 1–2, 172
Regulatory events, 152–153, 155–156
Regulatory genes, 8
Regulatory regions, 234
RegulonDB database, 54, 105, 110
  expression regulation, 117
  and protein gene regulation, 112
  and transcriptional units, 104
Regulons
  and Boolean networks, 207–208
  defined, 132
  prediction methods, 130–135
  purine biosynthesis, 133
  and transcriptional units, 116–117
Replication factors, 241
Reverse engineering, 213
RFC1-5, 241
Rhodopsin-like G-protein-coupled receptor (GPCR), 38
Ribosylation, 237
Ribulose uptake pathway, 266
RNA-binding motif, 38
RNA-polymerase (RNAP), 63–64, 106–117
RNP-1, 38
Rosetta Stone method, 225–228
RuleBase, 22, 24, 25
Rule-based modeling, 60–66

Saccharomyces cerevisiae. *See also* Yeast
  comparative analysis, 36–41
  DNA regulatory motifs, 139–140, 142
  mRNA expression, 131
  repeated domains, 41
Saccharomyces Genome Database (SGD), 33
SAGE, 237
Searches, 30–33, 55, 57–58
  RegulonDB, 117
  SQL, 73, 75
  SYSTER, 28
Semi-Thue system, 60
Sensitivity amplification, 193
Sequence Retrieval System (SRS), 57
Sequences. *See also* DNA sequences
  alignment, 260
  annotation, 20–25
  binding specificity (*see* Specificity)
  coding regions, 2
  databases, 49
  DNA, 5, 135–140
  domain repetition, 40
  P13K-Akt, 185
  regulatory network, 104, 106–117
  similarities, 229, 265
  specific, 87
Serum, growth factors in, 183
Sigma factor, 106
SignalP, 27, 37
Signal transduction pathways
  annotation, 185, 186, 188
  basic concept, 184
  as Boolean network, 204–209
  cascades, 193, 204–205, 208
  for chemotaxis, 280–282
  conflicting signals, 203
  and context, 208–209
  convergence, 187
  cooperativity, 193
  erythrocytes, 160–162

homology, 223
modularity, 190–191
perturbations, 189–190
simulation, 280–282
subcellular, 212–213
Similarities
explanations, 6–8
nonhomologous, 5
in phylogenetic profiles, 229
statistical significance, 28–29, 265
Simulation. *See also* Models
of cell (*see* E-CELL)
chemical kinetic, 164–167
E. coli chemotaxis, 281
essential data, 285–286
metabolic pathways, 268
of networks, 163–167
signal transduction, 191–192, 214
software for, 275–277
stiffness problem, 276
Simulators, 49–50, 58
Size, of proteins, 36–37
Sodium transport query, 30–32
Software, simulation, 275–277
Specificity
definition, 109
of DNA-binding protein
quantitative aspect, 87–89
weight matrix models, 89–97
RNAP transcription, 106–107, 110–117
to species, 263
Specificity score, 138–139
SPTR database, 26, 28
SQL. *See* Structured Query Language
SREBP proteins, 155, 159–160, 168
Stanford Microarray Database, 235
State space, 197, 198–200, 203, 214
Statistical analysis, 27, 31
Statistical significance
of similarities, 28–29
Sterol-regulated protease, 155–156, 160
Stiffness, 276
Structure-function relationships, 266
Structured Query Language (SQL), 73, 75
SWISS-PROT
annotation, 21–23, 24
and clustering, 28
and GO, 34
and protein functions, 238, 266
and proteome analysis, 26
controlled vocabulary, 33
database, 54
Synchronous congruence, 119
SYSTERS (SYSTEmatic Re-Searching), 28

T. pallidum, 266
Thermotoga maritima, 266
TIGR. (The Institute for Genomic Research), 261, 264
TIM barrels, 266–267
Tissue dynamics, 210
Tissues, morphogenesis, 160–163
TMHMM, 27, 37
Transcriptional units (TUs)
in bacteria, 116–117
boundary genes, 116, 121, 122, 124
Transcription factors
binding sites, 94, 98, 129
and Boolean networks, 208, 213
and cholesterol, 155, 159
and cis-regulators, 213
of erythrocyte differentiation, 162
Gene-Net depiction, 154, 155
for purine biosynthesis, 132
site recognition, 234
Transcription regulation
activators *vs.* suppressors, 114–115
in bacteria, 114
evolutionary origin, 105, 121
false positives, 106, 109, 111–112
and gene function, 114
initiation, 104, 106
Transcription Regulatory Regions Database (TRRD), 54
Transcriptomes, 2
TRANSFAC database, 54
Transition complexes, 193
Transmembrane proteins, 27
TRANSPATH, 49
TrEMBL
annotation, 21–23, 24
and clustering, 28

controlled vocabulary, 33
and GO, 34
and proteome analysis, 26
TRRD. *See* Transcription Regulatory Regions Database
Tus. *See* Transcriptional Units

Von Willebrand factor, 36

WD repeat, 38
Weight matrices
and DNA regulatory motifs, 139
of *E. coli* promoter, 107, 108–110, 113
expectation-maximization (EM), 96
Gibbs′ sampling, 96–97
greedy algorithms, 95
limitations, 110
network method, 97
in RegulonDB, 111
as specificity model, 90–93
Wings, 4, 5
WIT database, 54, 140–141, 231, 261

Yeast. *See also Saccharomyces cerevisiae*
and domain fusion, 227
in network, 241–242, 243
phylogenetic profile, 230
Yeast Protein Database (YPD), 135, 142