

Index

- 1D genome annotation 10, 34
- 1D/2D objects 455–458
- 2D genome annotation 10, 10, 13–15, 33, 34, 151
- 3D cell arrangements 479

- absolute chemical reaction rates 267
- abstract/specific knowledge 268
- adaptation dynamics 425–426
- adaptive laboratory evolution (ALE) 56, 422, 437
 - adaptation dynamics 425–426
 - adjustable constraints 471
 - applications 433–435
 - challenges 429–431
 - colonization evolution 434–435
 - complex adaptations 429, 433–437
 - D. melanogaster* 436–437
 - distal causation 422, 425, 470
 - E. coli* 425, 429, 433–435
 - endpoints 72, 354, 423, 425–428
 - experiments, performing 423
 - GEMs 72, 423, 428
 - genetic basis 424–427
 - genetic change types 425
 - genotype–phenotype relationship 422–423
 - hypothesis generation 428
 - molecular biology 427–428
 - nutrient adaptation 431–432
 - objective function 354, 428, 469
 - omics data sets 428
 - outcomes 427–431
 - pathogen adaption 435–436
 - perturbation 428
 - prediction failure 431
 - P. aeruginosa* 435
 - reviews 433, 437
 - selection pressure 428
- adenosine triphosphate *see* ATP
- adjacency matrices 180–182, 181
- adjustable constraints 256, 258, 289, 471
- ADOMETA 411, 413, 418–419
- aerobic growth state, sensitivity analysis 364–366
- Aeropyrum pernix* 418
- aggregate variables 474
- ALE studies *see* adaptive laboratory evolution
- algae *see* *Chlamydomonas reinhardtii*; *Synechocystis*
- allosteric constant 26
- alternative optimal solutions (AOS) 327
 - characterizing alternatives 327–330
 - E. coli* 340–341
 - enumerating alternatives 339–341
 - extreme pathways 329
 - flux variability analysis (FVA) 329
 - genome-scale studies 340–341
 - linear programming problems 304
- Alzheimer’s disease 244–247
- amino acids 275
 - Geobacter sulfurreducens* 93
 - regulon reconstruction 139–141
- anabolism, *E. coli* 54
- anti-infectives 80
- antibiotics 80
- AOS *see* alternative optimal solutions
- ATP (adenosine triphosphate) 18–19, 20
- ATP production 312–313, 315–320
 - aerobic production 315–319
 - anaerobic production 319–320
 - ATP balance 315, 319
 - flux maps 317, 319
 - linear programming 303
 - NADH aerobic production 321
 - NADH anaerobic production 325
 - objective functions 315, 344
 - optimal metabolic states 360
 - optimal yields 54, 313, 314
 - oxygen effects 361–362
 - phenotypic phases 360
 - PhPPs 371
 - reaction bi-linearity 268
 - see also* glycolysis pathway
- ATP F0F1 synthase 367
- ATP10A gene 396
- ATPase 282
- automated reconstruction 35–36, 47
- awareness (biological) 260, 472

- bacteria 72, 132, 279, 474–476
- baker’s yeast *see* *Saccharomyces cerevisiae*
- balance constraints/balances 279, 308, 461
- basis for vector spaces 167, 205
- basis vectors
 - computation 300
 - convex 167, 196, 214
 - non-orthonormal 195
 - rotating 240–241

- see also* orthonormal bases
- Bayesian networks 268, 417
- bi-level optimization procedures 444
- bi-linearity, biochemical reactions 193–196, 267–268
- bi-molecular association 154
- bias
 - biological network states 221
 - COBRA methods 250, 299–300, 470
 - disciplinary 457
 - network modules 127
 - network reconstructions 456, 457
 - objective function 353
 - pathway definitions 213
 - phylogenetic tree 75–77
 - SVD of **S** 202–203
 - topological properties 459
 - see also* randomized sampling
- bibliome 11
- bibliomic data 13, 15, 73, 126, 453
- big data, conversion to knowledge, (omic data analysis) 398, 401, 454 *see also* omics data; polyomic data
- BiGG (biochemical, genetic, and genomic) information 2, 8
- building 33–34
- conversion to mathematical format 12, 40, 40–42, 454, 458
- iJO1366 *E. coli* 65–67
- network reconstructions 13, 454
- S** matrix 40–42
- binary form of **S** 172–173
- binding sites of proteins 15, 118, 123
- biochemical network reconstructions 117, 133, 307
 - functional state combinatorics 268–270
 - integrating 128–130
 - pathway subsystems 417
 - protein structures 117–122
 - signaling networks 131–133
 - and statistical network models 398, 399
 - Tr/Tr networks 123–127
- bioengineering objectives 354
- biofuels 101
- bioinformatics 455
 - ADOMETA 418–419
 - draft reconstruction 457
 - gene functions 417
 - genome metastructures 135
 - integrative approach 3
 - reconstruction to model conversion 127
- biological attributes of **S** 152
- biological causation 11, 251, 252–253, 263, 461
 - vs. chemical 7, 254, 479
 - vs. physical 251
 - see also* distal causation; proximal causation
- biological components 3, 264–265
 - E. coli* GEMs 73
 - essentiality 73, 78, 378
 - measuring 453
 - part lists 13–14, 455–458
 - Principal Component Analysis (PCA) 240
 - self-assembly 260
 - vs. systems viewpoints 3, 264–265
 - see also* gene entries; metabolites; reactions
- biological constraints 260–262, 272–276, 293, 471
- biological discoveries, *E. coli* GEM 474–476
- biological functions 8, 468
- biological gap 410
- biological hierarchies *see* hierarchies, biological
- biological information theory 472
- biological networks
 - awareness 260
 - component self-assembly 260
 - higher-order properties 260
 - integrating reconstructions 128–130
 - multi-functionality 259
 - non-causality 259
 - redundancy 259
 - selection 260
 - self 260, 472, 480
 - systemic properties 270–271
- biological quantities, map of **S** 161–164
- biological systems
 - model-building 271
 - modularization 7
 - parameter diversity 255
 - time-dependency 255, 265
 - see also* hierarchies, biological; system boundaries; systems biology
- biology discipline 3, 272, 470–471
- biomass formation 344–349
 - bi-level optimization 444
 - biomass synthesis 346
 - C. reinhardtii* 102
 - driving cost 346
 - E. coli* 62, 346, 347, 444
 - input/output analysis 347
 - maintenance requirements 348
 - optimal yield 348, 364
 - phosphate/oxygen ratio 366
 - SVD 197–199
- biomass objective functions (BOFs) 68, 344, 345, 350–351, 355

512 INDEX

- biophysical perspective 478–479
- biosynthesis, drivers 346
- biosynthetic precursors 313–315
- biosynthetic prokaryotes 80–82
- bipopulations 252
- BLAST 416
- blocked reactions 410
- blue-green algae 80–82
- BNICE 411, 414–415
- BOFs *see* biomass objective functions
- Boltzmann, Ludwig 7, 255
- Boolean models 268
 - gene–protein–reactions 40, 105, 106, 132–133
 - M-models 294
 - signaling networks 133
 - transcriptional regulatory constraints 290–291, 474
- bottom-up reconstruction
 - CoSy biology 85
 - gap-filling 47
 - mathematics 10
 - network links, importance of 457
 - process descriptions 17, 33, 35–42, 49, 454
 - rationale 2, 275, 453
 - TRNs 138, 140
- bottom-up/top-down reconstruction 275, 453
 - TRNs 138
- bound states of PFK 26
- boundaries, system *see* system boundaries
- boundary fluxes 169
- bounds (boundary constraints) 308, 461
 - fluxes 202, 208–212, 292–293
 - mathematical representation 278, 279
 - narrowing alternative states 280
 - null space of **S** 214, 214, 300
 - theory vs. constraint-based thinking 271–272
 - see also* flux variability analysis; system boundaries
- Brenner, Sydney 145
- Buchnera aphidicola* 88
- cancer chemotherapy 107, 108, 403
- capacity constraints 282–285
 - E. coli* 282–284
 - energy-transducing membranes 283
 - flux variability analysis (FVA) 330
 - maximal reaction rates 285
 - membrane surface and metabolism 282
 - membrane surface and signaling 282–284
 - typical reaction rates 285
- carbon to oxygen uptake (C/O) ratio 363, 364
- carcinogenic process 262
- catabolism, *E. coli* 54
- Caulobacter crescentus* 417
- causal SNPs, mapping using co-sets 396
- causation 251–255
 - ALE studies 425
 - biological vs. physico-chemical 254, 479
 - dual 11, 251–253, 263, 343, 461
 - mathematical definitions 480
 - physics 251
 - see also* distal causation; proximal causation
- cell factory design 438–440, 443, 443–445, 446–450
- cells 307
 - 3D cell arrangements 479
 - constraint categories 272–273
 - constraints 277–278
 - evolution 12
 - GEM axioms 11–12
 - hierarchical 293
 - in silico* analog 5
 - interactions between cell parts 13–14
 - lifespan 264
 - maintenance 363, 364
 - molecular biology 478–479
 - multiple types 108–111
 - part lists 13–14, 455–458
 - phenotypic capacity 145
 - sensitivity analysis 363, 364
 - see also* biological components; objective functions
- charge balancing 160
- chemical attributes of **S** 151
- chemical causation 10
- chemical interaction networks 457
- chemical reactions *see* reactions
- chemical transformations 266–267
- chemistry
 - constraints from 11, 260–261, 285–289
 - GEM axioms 11
 - genetic code prediction 275
 - model-building approaches 271–272
 - molecular component links 266–267
 - S** as a data matrix 153–161
- chemotherapy 107, 108, 403
- ChIP-chip (or ChIP-exo) data 144
- Chlamydomonas reinhardtii* (green algae) 101–103
- chromosomal clustering 417–418
- citramalate pathway for isoleucine synthesis 94
- closed networks 170
- closed paths 288
- closed reaction systems 163, 205
- closed vector spaces 208–209
- clustering, chromosomal 417–418

- co-sets *see* correlated sets of reactions
- COBRA (constraint-based reconstruction and analysis) 12, 250, 272, 298
 - applying 13, 470
 - bias 250, 299–300, 470
 - biological hierarchy 478
 - brief history of 467–468
 - common misunderstandings 469–470
 - in silico* methods 299–300
 - methods 299–300
 - methods deployment 311
 - multiple possible solutions 470
 - objective functions 345, 355, 469, 470
 - optimization methods 298–299
 - scope 470
 - toolbox 300, 311
- co-factor molecules 154, 159, 268, 312–313, 449
 - see also* ATP production; NADH; NADPH
- colonization, evolution during 434–435
- column space of **S** 167–168
- communities (microbial) 88–90, 375–377, 478
- community systems (CoSy) biology 85
- competition (metabolic) 86
- complexity, biological phenomena 11
 - ALE studies 427, 429, 433–437
 - colonization 434–435
 - equivalent states 270
 - genome metastructures 134
 - intracellular environment 255
 - life cycles 429–431
 - processes 454
 - see also* bottom-up reconstruction; hierarchies, biological; modularization; subsystems
- components *see* biological components
- compound (node) adjacency matrix 181, 179–182
- compound maps of **S** 161, 162, 163
 - see also* nodes
- compounds vs. reactions 163, 300
- computational models *see* GEMs; *in silico* models; M-models; ME-models; model-building; quantitative models
- concentration variables 167, 461
 - constraint types 296
 - enzymes 284
 - fluxes and 165–166, 189, 208, 209
 - metabolites 296
 - SVD 189
 - water solvent capacity 284
- concentration vector 167
- confidence scores (reconstruction) 43
- connectivities 175–180, 182
 - connectivity distribution 177
 - connectivity matrix, **S** as 153, 161–165
 - connectivity number 173
- conservation relationships 151, 155, 167, 192
- conserved quantities 156
- constraint-based reconstruction and analysis *see* COBRA
- constraint-based thinking, vs. theory-based 271–272
- constraints 5, 277, 296, 297
 - adjustable 256, 258, 289, 471
 - bacteria 279
 - balances 279, 286, 308, 461
 - biological 260–262, 293, 471
 - biological consequences 272–276
 - biosynthetic networks 313
 - cellular 272–273, 275
 - from chemistry 11, 260–261, 285–289
 - coupling 291–295
 - dominant 210–211, 260–262, 281–282, 471
 - E. coli* 261, 277, 282–291
 - early concepts 258
 - energy 313
 - environmental 11, 251, 261–262, 273
 - evolutionary change 262, 275–276
 - fast subsystems 287
 - on flux values 201
 - flux variability analysis (FVA) 330
 - flux vector 210–212, 230
 - fluxomics 288–289
 - functional states 272–276
 - GEM axioms 11–12
 - general categories 273
 - genome-scale 277–278
 - H. pylori* 261
 - hard 256, 260, 273, 289, 471
 - imposing 279–282
 - in silico* models 256, 279
 - integer cut 340
 - mass conservation 12, 286–287
 - mathematical representation 279–282
 - membrane surfaces 282–284
 - non-adjustable 256, 260, 273, 289, 471
 - omics data as 250, 403–405
 - physico-chemical 12, 260–261, 285–289
 - polyomic data 296
 - quantitative models 256, 279
 - quasi-steady states 287
 - redundant 210–211, 236, 281–282, 454
 - self-imposed 262
 - simultaneous satisfaction 296
 - solution spaces 230, 296
 - spatial 261

514 INDEX

- stating 279–282, 300
- steady-state assumption 286
- steady-state flux vector 210–212, 230
- stoichiometry of chemical reactions 314
- successive imposition 279–281
- vs. theory 271–272
- thermodynamics 287–288
- transcriptional regulatory 290, 291
- see also* capacity constraints; regulatory constraints
- content 398
- see also* omics data
- conversion processes 77
 - BiGG reconstructions to mathematical format 12, 40, 40–42, 307, 454, 458
 - genome-scale reconstruction to GEM 308, 308–311
 - omics data to BiGG reconstruction 398, 401, 454
 - Tr/Tr networks 126–127
- convex bases 167, 196, 209, 214
- correlated sets of reactions (co-sets) 80, 127, 227, 228, 396
- correlated subsets of reactions 222, 225–227, 228
- CoSy (community systems) biology 85
- coupled GEMs 375–377
- coupling constraints 291–295
 - biological hierarchy 291
 - coupling levels 292–293
 - formulation of 293–295
- Crick, Francis 73
- crosstalk 222, 229
- curation 37–40, 457
 - genome metastructures 135–136
 - iJO1366 68
 - Recon 1 104, 105
 - time scales 43
 - Tr/Tr networks 126
- cyanobacteria 78, 80–82
- cystic fibrosis 435
- Danchin, A. 273
- Darwin, Charles 7
- data 8, 144–145, 296
 - conversion to knowledge 401, 454
 - genome metastructures 144–145
 - high-dimensional 237–241, 474
 - high-throughput 77, 106–107, 401
 - metabolomics 238–239
 - Recon 1 106–107
 - S** as a data matrix 153, 161
 - see also* bibliomic data; BiGG information; databases; omics data; polyomic data; randomized sampling
- databases 8
 - OMIM database 108
 - ORENZA database 411
 - Protein Data Bank (PDB) 123
 - see also* BiGG information; knowledge bases
- dead-end metabolites 47, 410
- decompositions
 - system 472
 - variable pools 480
 - see also* singular value decomposition
- degenerate solutions 304
- dFBA 469
- diabetes 111–114
- DIET (direct interspecies electron transfer) 375–377
- DIP metabolite 417–418
- directionally coupled reaction pairs 227
- directions (column space) 199–201
- disciplinary bias 457, 478–479
- distal causation 5, 11, 342, 355, 422
 - ALE studies 470
 - biological attributes of **S** 152
 - biological networks 270–271, 461
 - biomass objective function 350–351
 - biomass production 344–349
 - environmental parameters 377
 - objective functions 12, 342–343, 353–354, 469
 - vs. proximal causation 343
 - studying objective functions 353
 - types of objective functions 343, 344
 - see also* evolutionary change
- diversity 252, 342
- DMMM (dynamic multi-species metabolic modeling) 86
- DNA molecule 15, 273, 275
- DNA sequence 10, 33
 - see also* genome annotation
- DNA structure 474
- dominant constraints 210–211, 260–262, 281–282, 471
- double gene knock-outs 390–395
- draft reconstruction 35–36, 457
- drivers (row space) 201–202
- Drosophila melanogaster* (fruit flies) 436–437
- drug action mechanisms 123
- drug effects 107
- drug targets
 - binding sites of drug molecules 123
 - cancer chemotherapy 107, 108, 403

- discovering 378, 403
- essential metabolites 78
- groups 80
- multiple 80
- pathogen networks 79
- Recon 1 107, 108
- Recon 2 115
- drug-treated states, simulations 108
- dual causation 11, 251, 263, 461, 462
 - see also* distal causation; proximal causation
- dual-species models 77
- dynamic flux balance analysis (dFBA) 54, 469
- dynamic flux distributions 167
- dynamic multi-species metabolic modeling (DMMM) 86
- dynamic states 201
 - vs. steady 469
- E-matrix (Expression matrix) 127, 128, 130, 291
- educational values 464, 465
- eigenvector rotation 240
- elemental balancing 160
- elemental matrix 155–156
- emergent properties *see* network properties
- endpoints 72, 354, 423, 425, 426–427, 428
- energy balance analysis 309
- energy constraints 313
- energy needs for growth 335
- energy-transducing membranes 283, 363, 364, 366
- engineering discipline 11
 - fast subsystems 287
 - model-building approaches 77, 271–272
 - perception of solution 470
 - steady states 469
- environmental constraints 11, 251, 261–262, 273
- environmental parameters 12, 359, 377
 - distal causation 377
 - oxygen effects 361, 362–364
 - phenotypic phase planes 368–371
 - robustness analysis 359–360, 366–368
 - sensitivity analysis 364–366
 - varying a single parameter 359
 - varying two parameters 368–371
- enzymes
 - concentrations 284
 - coupling levels 292–293
 - efficiency 117–119
 - kinetics 285
 - new reactions 420
 - phosphofructokinase 25–29
 - promiscuity 259, 406
 - reaction rates 285
 - robustness 78, 366
 - specificity 117–119, 406
 - and transporters 18–19
- epigenotype *in silico* models 395–396, 396
- epistasis 378, 391
- equivalent functional states 270, 327, 341
 - AOS extreme pathways 336–339
 - enumerating AOS 339–341
 - flux variability analysis (FVA) 330–336
 - network objectives 327–330
- equivalent optimal solutions 259, 304, 309, 471
- equivalent reaction sets 335
- Escherichia coli* 29–31, 51–65
 - 3D biophysical view 278
 - aerobic ATP production 315–319
 - ALE studies 425, 429, 433, 435
 - anaerobic ATP production 319–320
 - AOS 340–341
 - B. aphidicola* 88
 - biomass production 62, 346, 347, 444
 - capacity constraints 282, 284
 - catabolism 54
 - cell wall/co-factor biosynthesis 55
 - co-sets 227
 - double gene knock-out 391–395
 - environmental constraints 261
 - environmental parameters 362–364
 - enzyme gaps 410
 - essential components 73
 - extreme pathways 337
 - flux coupling finder (FCF) 228
 - functional states 267, 268, 270
 - gap-filling 53, 70, 412–414
 - GEM applications 70–73
 - GEMPRO model 121
 - genetic parameters 379–383, 391–395
 - genome-scale constraints 277
 - iAF1260* 57–59, 352, 393, 410
 - iJE660* 55–56, 352, 383–384
 - iJO1366* 59, 65–70, 121
 - iJR904* 56–57, 352, 419
 - K-12 MG1655 strain 50, 59–64, 410, 431
 - metabolic engineering 72
 - metabolism 50–51, 73
 - metabolite coupling 182
 - MOMA analysis 387
 - new reactions 420
 - objective functions 352, 353, 362–364
 - optimal network performance 312–315
 - P/O ratio 315, 366
 - pathway vectors, co-sets 227

516 INDEX

- pathways 219, 312
- pre-genome era 53–55
- Project K 73
- randomized sampling 237, 238
- S** matrix 175, 176
- Shigella* strains 59–64
- signaling network 132
- single gene knock-outs 31, 31–32, 379–383
- species reconstruction 59–64
- stimulons 140–141
- SVD 196–199
- tangle of cycles 269
- transcriptional regulatory constraints 291
- variable fluxes 332–335, 340–341
- essentiality of components 73, 78, 378
- Euclidean distance 299, 387
- eukaryotes 96, 116
 - available reconstructions 45
 - C. reinhardtii* 101–103
 - genome-scale reconstruction 96
 - H. sapiens* 103–116
 - S. cerevisiae* 96–101
 - signaling networks 132–133
- evaluations 8, 42, 457
- evolutionary change
 - ALE studies 434–435
 - biological networks 270–271
 - during colonization 434–435
 - constraint adoption 275–276
 - hierarchy formation 275–276
 - objective function 355
 - and optimization 5
 - regulatory constraints 262
 - timeline 275
 - see also* adaptive laboratory evolution; distal causation; selection pressures
- exchange fluxes 215
- exometabolites 115
- exometabolomic (ExMe) data 238–239
- ExPa/ElMo 309
- Expression matrix (E-matrix) 127, 128, 130, 291
- extreme pathways
 - classifying 215–216
 - computing 214–215, 217–219
 - exchange fluxes 215
 - glycolysis 217
 - and optimal states 336–339
 - red blood cell metabolism 338
 - simplest linear basis 217
 - simultaneous functions 338
 - two networks 338–339
 - type II elimination 216
 - type III elimination 216
 - types 215
- extreme points 211, 303, 329, 340
- extremophile metabolism 374–375
- failure, of predictions 409, 431
- false-negative predictions 69–70, 384, 409, 412–414, 416
- false-positive predictions 70, 108, 384, 409, 415–416
- FBA *see* flux balance analysis
- FCF (flux coupling finder) 227
- Fellowship for Interpretation of Genomes (FIG) 417
- Fick's law 251, 473
- flux balance analysis (FBA) 309, 470
 - COBRA 250
 - COBRA toolbox 301
 - dynamic (dFBA) 54, 469
 - gene knock-outs 379
 - GENREs 77
 - linear programming 300
 - Tr/Tr networks 125
- flux balances 286, 330
- flux capacity constraints 296
- flux confidence interval 309
- flux coupling 228, 309
- flux coupling finder (FCF) 227
- flux maps 460
 - ATP production 317, 319
 - extreme pathways 215, 214–215, 223
 - NADH 322–324
 - parsimonious 405
 - and pathway length 222–224
- flux span 331
- flux splits 211, 233–234
- flux variability analysis (FVA) 309, 332, 330–335, 336
 - alternative optimal states 329
 - concept 330–332
 - E. coli* 332–335, 340–341
 - energy needs for growth 335
 - equivalent reaction sets 335
 - genome-scale results 335–336
 - input/output analysis 335
 - metabolic demands 335
 - oxygen uptake rate, effects on growth rate 364
- flux variability spectrum 331
- flux variable dependency 331–332
- flux vector 185, 344
- fluxes and concentrations 189, 195, 208, 209
- fluxes, boundary 169

- fluxomics 15, 288–289
 - COBRA toolbox 301
 - computations 288–289
 - pairwise omic data integration 146
- fruit flies (*Drosophila melanogaster*) 436–437
- fumarate hydratase (FH) 393
- functional genome annotation 135
- functional role set 417
- functional states 172, 264, 276, 459–461
 - alternative 281
 - awareness 260, 472
 - biological functions 8, 267–271
 - components vs. systems 264–265
 - computing 454, 459–462
 - constraining 271–272
 - finding 300–301
 - gene and genome correlation 270
 - link properties 266–269
 - network capabilities 459–462
 - of a single network 268–270
- fundamental subspaces of **S** 166–167, 171, 184
 - column space 167–168
 - contents 167
 - dimensions 166–167
 - drivers and directions 201–202
 - left null space 167–168
 - linear reversible reaction 192
 - row space 167, 168
 - S** as a linear transformation 166–167
 - SVD 184–188, 202, 203
 - see also* null space of **S**; orthonormal bases
- futile cycle 216
- futile phases 370
- future work 132–133, 145, 473–480
- FVA *see* flux variability analysis
- gain-of-function (GOF) 119
- GAM (growth-associated maintenance) 348
- gap-filling 47, 250, 407–411, 420–421
 - algorithms 411–416
 - COBRA toolbox 301
 - dead-end metabolites 47, 410
 - E. coli* iJO1366 53, 70, 412–414
 - false-positive predictions 70, 409, 416
 - gap reasons 410
 - gap types 409–410
 - gene functions 416–420
 - information requirements 407
 - initial curation 68
 - orphan reactions 410–411
 - prediction failure 407–409, 431
 - prediction methods 411, 413
 - reaction gaps 409–410, 412–414
 - reactions 412–416
- GapFill/GapFind 309, 411, 413, 416
- GEM-PRO model 121
- GEMs (genome-scale models) 31, 32, 307, 357, 472–473
 - ALE studies 423, 428
 - applications 70–73, 449, 462–463, 474, 476
 - axioms 11–12
 - building quality models 43, 306–308, 458
 - capabilities of 325
 - cell factory design 446–450
 - COBRA 298, 306–311
 - conversion from reconstruction 308–311
 - coupled 375–377
 - current position 31
 - E. coli* applications 70–73
 - expanding the scope of 474
 - formation process 308
 - gene interactions 392–393
 - Geobacter metallireducens* 91, 92, 375–377
 - Geobacter sulfurreducens* 88–94, 96–101
 - guided strain design 448
 - history 96, 468
 - limitations 474, 476
 - metabolic engineering 449
 - missing information 407
 - missing regulation 474
 - need for 8
 - omic data 399–401
 - parameter values 277
 - parameters, number of 470
 - photosynthesis 82
 - properties, global view 311
 - S. cerevisiae* 96–101
 - scale-independent applications 325
 - self/non-self distinction 472
 - strengths 474, 476
 - system boundaries 68, 308, 457
 - time-dependency 474
 - see also in silico* models
- gene deletion 250, 300, 309, 383, 384
 - see also* gene knock-outs; single gene knock-outs
- gene dosage 378, 395–397
- gene essentiality analysis 78
- gene expression 296, 395, 472
- gene functions, predicting 407, 416–420
- gene imprinting 395–396
- gene interactions 391–393
- gene knock-outs 378–379, 390–395, 468
 - see also* gene deletion; single gene knock-outs

518 INDEX

- gene perturbation 250
- gene–protein–reactions (GPRs) 31–32
 - associations 36, 37, 39–41
 - BiGG information conversion to mathematical format 40
 - data types 36, 37
 - draft reconstruction 36, 37
 - enzyme properties 117
 - gene knock-outs 31, 378, 379
 - iJR904 56
 - M-models 294
 - new reactions 420
 - Recon 1 105, 106
 - reconstruction process 29
 - strain designs 446
- gene/genome correlation 270
- generalist phenotype 259
- genetic change types 425
- genetic circuits 3, 4
- genetic code prediction 275
- genetic parameters 378, 397
 - double gene knock-outs 390–395
 - E. coli* 379–383, 391–395
 - epigenotype *in silico* models 396
 - gene dosage 395–397
 - gene imprinting 396
 - genome-scale studies 383–386, 392–395
 - single gene knock-outs 378–379
 - studying non-lethal gene KOs 386–390
- genetic/genomic attributes of **S** 152
- genome annotation 12, 468
 - 1D annotations 10, 34
 - 2D annotations 10, 10, 13–15, 33, 34, 151
 - functional 135
 - GEM axioms 11
 - Haemophilus influenzae* 50
 - limitations 134
 - metastructures 135
 - refinement algorithm 309
 - subsystems approach to 417–418
- genome metastructures *see* metastructures of genomes
- genome re-sequencing, ALE outcomes 427
- genome-scale, redundancy 259
- genome-scale matrices
 - connectivities 175–179, 182
 - H. pylori* 224, 225
 - participation 182, 225
 - pathway length 224
- genome-scale metabolic models (M-models) 294
- genome-scale metabolic network reconstructions
 - 8, 29–32, 35–42, 49, 457
 - applications 48, 77
 - categories 77
 - COBRA 298
 - conversion to models 308–311
 - E. coli* applications 70–73
 - eukaryotes 96
 - GEM axioms 11
 - GENREs 77
 - H. influenzae* 50
 - K-12 MG1655 *E. coli* strain 50
 - naming conventions 55
 - prokaryotes 77
 - S. cerevisiae* 96
 - Synechocystis* 82
 - see also* biochemical network reconstructions; biological networks; bottom-up reconstruction; curation; *Escherichia coli*; gap-filling; genome-scale metabolic network reconstructions; metabolic network reconstructions; modularization; Recon 1; system boundaries; workflows
- genome-scale models, generations 257–259, 258
- genome-scale perspectives
 - constraints 277–278
 - double gene knock-outs 392–395
 - extreme pathways 337–339
 - genetic parameters 383–386, 392–395
 - integration, *T. maritima* 120, 130
 - phenotypic phase planes 371–377
 - single gene knock-outs 383–386
 - stoichiometric models 293
- genome-scale science 3–7, 478–479
- genomics 15, 146
- genotype–phenotype relationships 1, 2, 293
 - ALE studies 422–423
 - construction 1, 2, 32, 73, 254
 - current position 31
 - environment 12
 - GEM axioms 11, 12
 - see also* GEMs; *in silico* models; quantitative models
- GENREs (genome-scale reconstructions) 77
- Geobacter metallireducens* 91, 92, 375–377
- Geobacter sulfurreducens* 86, 88–90
 - acetate activation 91
 - acetate uptake 90
 - amino acid metabolism 93
 - characterization history 89
 - DIET 375–377
 - GEMs 88–94
 - genome-scale science 90
 - metabolic challenges 90

- S matrix 173
 - workflow 90, 91
- Gibbs free energy 287–288
- global orphan reactions 411
- glutamate 9
- glycolysis pathway 18–22, 19, 20
 - aerobic ATP production 315–319
 - compound maps of **S** 164
 - compounds 19
 - computation 217
 - coupling to pentose phosphate pathway 22
 - enzymes and transporters 18–19
 - NADH aerobic production 320–325
 - nodes and links 18–19, 181
 - and oxygen 362–364
 - participation and connectivities 181
 - PFK enzyme 25–29
 - robustness analysis 366
 - S** matrix 17–20, 25, 158–161
 - see also* ATP production
- Gödel's completeness theorem 275
- GOF (gain-of-function) mutations 119
- governing constraints 210–211, 260–262, 281–282, 471
- GPRs *see* gene–protein–reactions
- green algae 101
 - see also* *Chlamydomonas reinhardtii*
- GrowMatch 385, 411, 413, 416
- Growth-associated maintenance (GAM) 348
- growth-coupling 444–446
- growth properties, gene knock-out strains 383–386
- growth rates 344
 - extremophile 374–375
 - glucose and oxygen 362–364
 - lethality 380–383
 - model-driven discovery 409
 - N. pharaonis* 374–375
 - oxygen uptake rate 362–364
 - PhPPs 371
 - single gene knock-outs 383, 379–383
 - see also* adaptive laboratory evolution; biomass formation
- Haemophilus influenzae*
 - extreme pathways 338–339
 - input/output states 229
 - PhPP 371–374, 374
 - reconstructions 50
 - single gene knock-outs 383
- Haig's parental conflict theory 395
- hard constraints 260, 273, 289, 471
- Helicobacter pylori*
 - environmental constraints 261
 - extreme pathways 338–339
 - flux coupling finder (FCF) 228
 - input/output states 229
 - metabolite coupling 182
 - pathway length 225
- HGP *see* Human Genome Project
- hierarchies, biological 273–275
 - cells 293
 - COBRA methods 478
 - coupling constraints 291–293
 - evolutionary change 275–276
 - multi-scale analysis framework 472, 479–480
 - nature of living systems 254
 - systems biology 7
- high-dimensional data sets 237–241, 474
- high-dimensional model-building, iterative
 - nature 407–409
- high-throughput (HT)
 - analytical chemistry 3
 - data 77, 106–107, 401
 - methods 15
 - Recon 1 106–107
 - technologies 3, 8, 13–14, 453, 455
 - see also* BiGG information; bioinformatics; omics data; polyomic data
- higher-level vs. lower-level functions 273–276
- hit fraction 237
- hitchhiker mutations 260
- hole filler program 416
- Homo sapiens* 103–116
 - brain 108–111
 - cardiac mitochondrion, metabolite coupling 182
 - cells and tissues 80, 264, 478
 - combinatorics 268
 - epigenotype *in silico* models 396
 - GEM expansion 478
 - health and disease 104
 - host–pathogen models 80
 - multi-cell reconstruction 108–111
 - multiple organ models 111–114
 - PFK enzyme 25–29
 - protein multi-functionality 282
 - Recon 1 106–108
 - Recon 2 114–116
 - Recon X 116
 - reconstruction jamboree 98
 - renewal rates 264
- homologous genes 107
- hot spots 121, 122

520 INDEX

house mouse (*Mus musculus*) 114, 386HT *see* high-throughput

Human Genome Project (HGP) 103

Hybridoma 352

hypothesis generation 47, 77, 407, 428

see also ADOMETA; gap-filling*iAF1260*, for *E. coli* 57–59, 352, 395, 410

IEMs (inborn errors of metabolism) 116

iJE660, for *E. coli* 55–56, 352, 383, 384*iJN678*, for *Synechocystis* 82*iJO1366*, for *E. coli* 59, 65–70, 121*iJR904*, for *E. coli* 56–57, 352, 419*iMM1415*, for *M. musculus* 114

imprinted genes 378

in silico models (computational models) 3, 8,
308–311

applications 454

biological systems 271

cell analogs 5

gene interaction 390–395

implementing constraints 279

mathematical models 68, 454

metabolites in 68

model-building approaches 67–68, 255–260,
271–272, 407–409, 472–473

naming convention 55

see also COBRA; GEMs*in vivo* gene essentiality data 87

inactivating gene products 300

inborn errors of metabolism (IEMs) 116

indels 425

inference-based analysis (top-down) 398, 399,
454*see also* bottom-up reconstructioninformatic attributes of **S** 151

information theory, biological 472

input/output

analyses 335, 347

feasibility array (IOFA) 228

relationships 222

signatures 228, 229

systems boundary 19, 68

integer cut constraints 340

integrated metabolic/network models 293–295

interactions between cell parts 13–14

interspecies interaction 474, 476

intracellular environment 255, 479

IOFA (input/output feasibility array) 228

iRC1080, for *C. reinhardtii* 101

isocitrate dehydrogenase (ICDH) 119

isocline 370

isoleucine 404–405

isotopomer distribution vector (IDV) 288

isotopomer models 288–289

iterative workflows *see* workflows

Jacob, F. 273

K-12 MG1655 *E. coli* strain 50, 59–64, 410, 431k-bases *see* knowledge bases

kinetics 199, 268, 268, 296

enzymes 285

parameters 296, 474

reaction rates 285

Kirchoff's circuit laws 286, 288

knowledge-based analysis 398, 399

knowledge bases (k-bases) 8, 13, 32–34, 49

see also BiGG information

knowledge gaps 410

knowledge spectrum 268

Lactobacillus plantarum 335left null space of **S** 167–168

left singular vectors 187

Leishmania major 367

lethality

non-lethal gene knock-out 386–390

single gene knock-outs 380–383

synthetic 378, 391–395

leucine responsive protein (Lrp) regulon,

reconstruction 139–141

life sciences 11–12

line of optimality 369, 371, 376, 424

linear algebra 280, 459

linear bases 167

convex vs. 208, 209

extreme pathways 217

non-negative 207–208

null space of **S** 205–208, 220

simple 205, 207

see also orthonormal bases

linear dependency 166

linear maps 165

linear programming (LP) 298, 299, 302–306

ATP production 315

finding functional states 301

objective function studies 352

linear reaction networks 205–207

linear reversible reaction 191

linear spaces 209

links in a network (reactions) 17, 267, 268

adjacency matrices 181

connectivities 179–180, 182

- determining 457
- glycolysis pathway 158
- S matrix 161
- Listeria monocytogenes* 404
- local orphan reactions 411
- location analysis (binding sites of proteins) 15
- LOF (loss-of-function) mutations 119
- loops, eliminating 288
- loss-of-function (LOF) mutations 119
- LP *see* linear programming
- Lrp (leucine responsive protein) regulon 139–141
- M-models (genome-scale metabolic models) 294
- macrophage activation, metabolic junctions in 405–406
- MAGE method 422
- maintenance energy 363, 364
- mammals, Recon 1 mapping 114
- Manhattan norm 344
- manual curation *see* curation
- mass conservation 12, 260, 286–287
- Mass Distribution Vector (MDV) 289
- mass spectrometry (MS) 288
- mathematical attributes of S 152
- mathematical model formulation 13, 472–473
 - constraints 279
 - formulating definitions 480
 - in silico* model formulation 68
 - see also* COBRA; conversion processes; GEMs; quantitative models
- MATLAB 300
- ME-models (metabolism and expression models) 130, 131, 294, 474
- mechanistic approach 453
 - genotype–phenotype relationship construction 11–12, 32, 73, 254
 - see also* GEMs; *in silico* models; quantitative models
- membrane surface capacity
 - metabolism constraints 282
 - signaling 282–284
 - water solvent capacity 284
- Mendel, Gregor 1
- metabolic engineering
 - COBRA toolbox 301
 - E. coli* GEMs 72
 - GEM applications 449, 474, 476
 - GENREs 77
 - robustness analysis 367–368
 - three-phase history 300, 440
 - see also* strain design; synthetic biology
- metabolic models 293–295
- metabolic network reconstructions 17, 32, 33, 49, 453
 - applications 45–48
 - automated reconstruction 35–36, 47
 - available reconstructions 45
 - bottom-up process 17, 33, 35–42, 49, 454
 - bottom-up rationale 2, 275, 453
 - COBRA toolbox 301
 - common issues 38
 - conceptual foundations 457–458
 - confidence scores 43
 - conversion to mathematical format 40–42, 454
 - DNA comparison 33
 - draft reconstruction 35–36, 457
 - E. coli* role 51–65
 - evaluation 42, 457
 - four steps 8, 35–42, 49, 308, 457
 - GEM axioms 11–12
 - H. influenzae* 50
 - iJN678, for *Synechocystis* 82
 - initial studies 467
 - integrating 128–130
 - integrating protein synthesis 130
 - iRC1080, for *C. reinhardtii* 101
 - iterative nature 44, 49, 51–65
 - jamborees 34, 98–99
 - limits of 471
 - Lrp regulon reconstruction 139, 141, 453
 - microbial communities 478
 - modularization 22–25, 27, 127, 138, 479, 480
 - multi-cell 108–111
 - multi-strain 478
 - naming conventions 55
 - operons 138–139
 - P. putida* 43
 - pathogens 78–81
 - quality procedures 43, 306–308, 458
 - RAVEN (reconstruction tool) 98
 - regulons 139, 140
 - resources 43–45
 - S. cerevisiae* (yeast) 100, 96–101
 - S. typhimurium* LT2, 98, 99
 - standard operating procedure 43–45
 - stimulons 140–141
 - system boundaries 19, 169–170, 471
 - T. maritima* 130
 - top-down 454
 - unbiased 456
 - validation 42, 457
 - see also* bottom-up reconstruction; bottom-up/top-down reconstruction; curation; *Escherichia coli*; eukaryotes; gap-filling;

522 INDEX

- gene–protein–reactions; genome-scale
metabolic network reconstructions;
prokaryotes; Recon X; Recon 1; workflows
- metabolism and expression models *see*
ME-models
- metabolites 284
 - coupling 182
 - dead-end 47, 410
 - E. coli* 182
 - elemental matrix 155–156
 - essentiality 78, 309
 - in silico* model formulation 68
 - non-native metabolite production 449
 - objective function 344
 - reporter 402
 - root no-consumption 410
 - root no-production 410
 - S matrix 167
 - turnover time 264–265
- metabolomics 15, 146
- MetaCyc 47, 416
- metastructures of genomes 134–138, 148, 478, 479
 - annotation types 135
 - mathematical descriptions 145–148
 - polyomic data integration 144–145
 - refactoring DNA 142–144
 - transcriptional regulatory networks 138–142
- Michaelis–Menten rate law 287
- microbial communities 86–88, 478
- MILP (mixed integer linear programming) 299,
299, 339–340, 416, 446
- miniaturization 453, 455
- minimization of metabolic adjustment (MOMA)
309, 386–390
- MinSpan 217–219
 - see also* gap-filling
- mixed integer linear programming *see* MILP
- model-building 67–68, 77, 255–260, 271–272,
407–409, 472–473
- model-driven design 438, 450
 - GEMs and cell factory design 446–450
 - GEMs and strain design algorithms 444–446
 - historical background 438–444
- model refinement, COBRA methods 250, 300
- Model SEED 47
- modularization 7, 479–480
 - modular units (MUs) 138
 - module-by-module reconstruction 22–25, 27
 - Tr/Tr networks 127
- molecular biology 1, 3, 73
 - ALE outcomes 427, 427–428
 - and cell biology 478–479
 - molecular crowding constraints 296
 - molecular machines 3
 - molecular systems biology 464
 - see also* systems biology
 - MOMA (minimization of metabolic adjustment)
309, 386–390
 - Monte Carlo sampling 234, 242, 250, 301
 - mouse studies (*Mus musculus*) 114, 386
 - mRNA molecules, turnover time 264–265
 - multi-cell reconstructions 108–111
 - multi-omic data 8, 375, 377
 - see also* polyomic data
 - multi-organ models 478
 - multi-scale analyses 34, 472, 479–480
 - multi-species relationships 77
 - multi-strain reconstruction 478
 - multidisciplinary approach 464
 - multiple targets of a single compound 80
 - multiple tissue types 104–108, 111–114
 - multiplexing 453, 455
 - Mus musculus* (house mouse) 114, 386
 - mutations
 - false positives/false negatives 416
 - gene deletion 381
 - GOF/LOF 119
 - hitchhiker 260
 - mutant growth 381
 - selection pressures 282
 - systems biology 252, 432
 - see also* adaptive laboratory evolution
 - mutualism 86
 - Mycobacterium tuberculosis* 87, 243, 245, 417
 - Mycoplasma pneumoniae* 284
- NADH 18–19, 312–313, 320–325
 - aerobic production 320–324
 - anaerobic production of 324–325
 - linear programming 303
 - NADTRHD 387
 - optimal yields 54, 313, 314
 - P/O ratio (phosphate/oxygen ratio) 366
 - Recon 1 393
- NADPH 312–313
 - aerobic production 320
 - alleviating the ATP constraint 323–324
 - interpretation of optimal solution 322–323
 - NADTRHD 387
 - optimal yields 54, 313, 314
- NADTRHD (transhydrogenase) 387
- naming conventions
 - flux vector/right singular vectors 185

- genome-scale network reconstruction 55
- Natronomonas pharaonis* 374–375
- natural selection 273
- Nernst equations 255
- net analysis 309
- net reactions 193
- network analysis approaches 398
- network gaps *see* gap-filling [reactions]
- network modules, Tr/Tr networks 127
- network perspective 478, 479
- network properties 184, 202–203, 249
 - determining capabilities 312, 325–326
 - directions (column space) 199–201
 - drivers (row space) 201–202
 - SVD 184–188
 - see also* network topology
- network reconstructions 8, 117, 133, 307
 - functional state combinatorics 268–270
 - integrating 128–130
 - jamborees 34, 98–99
 - protein structures 117–122
 - signaling networks 11, 12, 131–132, 133, 457
 - and statistical network models 398, 399
 - subsystems 417
 - teaching about 455–458
 - Tr/Tr networks 123–127
 - transcriptional regulatory networks (TRNs) 12
 - see also* BiGG information; bottom-up reconstruction; bottom-up/top-down reconstruction; *Escherichia coli*; gap-filling; genome-scale reconstructions; metabolic network reconstructions; Recon 1; Recon X; workflows
- network topology 12, 172, 182, 459
 - binary form of **S** 172–173
 - column shifting operators 174
 - compound connectivity 173
 - connectivities in genome-scale matrices 175–179
 - groups of drug targets 80
 - node connectivity and network states 178
 - participation and connectivity 173–180, 182
 - reaction participation number 173
 - rearranging the stoichiometric matrix 174–175
 - row shifting operators 174
 - S** as a connectivity matrix 153, 161–165
 - S** is a sparse matrix 172
 - Tr/Tr networks 127
- NGAM (non-growth associated maintenance) 348
- nitrogen
 - H. influenzae* requirements 371–374
 - regulon reconstruction 139–141
- NLP (non-linear programming) 299, 299
- NMR (nuclear magnetic resonance) 288
- node (compound) maps 169
 - ATP production 317, 319
 - NADH 322–324
 - redox tangle 269
- nodes (compounds) 11, 17
 - adjacency matrices 181
 - connectivity distribution 177
 - connectivity number 173
 - glycolysis pathway 158
 - high connectivity 175, 178
 - Kirchoff's circuit laws 286
 - linearity of maps 165
 - manual curation 457
 - pathway reconstruction 18–19
 - S** matrix 161
 - simple network 172
- non-adjustable (hard) constraints 256, 260, 273, 289, 471
- non-growth associated maintenance (NGAM) 348
- non-lethal gene knock-out 386–390
- non-linear maps 165
- non-linear programming (NLP) 299, 299
- non-native metabolite, production of a 449
- non-orthonormal basis vectors 195
- non-self (biology) 472, 480
- nuclear magnetic resonance (NMR) 204, 288
- null space of **S** 204, 214
 - basis vectors 204
 - bounds 214, 214, 300
 - contents 167
 - flux vector 168
 - internal and external reactions 205
 - left null space 167–168
 - linear basis 205, 207
 - linear v. convex spaces 209
 - non-negative linear basis 207–208
 - pathways 204
 - pools and pathways 153
 - spanning with pathway vectors 204–205
 - terminology 464
- nutrient uptake, objective function 344
- nutritional supplements 101
- obesity 113
- objective functions 250, 253, 301, 342, 355, 472
 - ALE 354, 428, 469

524 INDEX

- ATP production 315, 344
- bias 353
- biomass formation 68, 344, 345, 350–351
- COBRA 345, 355, 469, 470
- degenerate solutions 304
- distal causation 12, 342–343, 353–354, 469
- E. coli* 352, 353, 362–364
- evolutionary change 355
- GEM axioms 12
- general linear function 301
- linear programming 298, 301, 352
- M. tuberculosis* 87
- mathematical representations 301, 343
- metabolites 344
- MILP 299
- natural environments 354
- network properties 312
- non-linear programming 299
- nutrient uptake 344
- and optimization methods 298–300
- philosophical aspects 354
- PhPP 368, 370
- physico-chemical properties 344
- quadratic programming (QP) 298
- redox potential 320
- reduced costs 305
- Saccharomyces cerevisiae* 352, 353
- selection process 12, 342, 355
- shadow prices 305
- studying 353
- time-dependency 431
- types of 343–344
- unbiased searches 353
- unbounded solutions 305
- objective tilting 446
- objectives of biological networks 271
- ObjFind (objective function finder) 309
- omics data 3, 8, 12, 13, 398, 406
 - 2005–2009 468
 - ALE studies 428
 - cancer chemotherapy 107, 108, 403
 - as constraints 250, 403–405
 - contextualization 398–401
 - data types 127, 479
 - GEM predictions 405–406
 - GEMs 399–400
 - Geobacter sulfurreducens* 93
 - integration 144, 146
 - mapping and network topology 402–403
 - network models 398
 - omic data analysis 398, 401, 454
 - polyomic integration 375–377
 - Recon 1 107
 - Tr/Tr network mapping 127
 - see also* BiGG information; databases; high-throughput; polyomic data
- OMIM (Online Mendelian Inheritance of Man)
 - database 108
- OMNI 411, 415–416
- open networks 170
- open reaction maps 163
- open reading frames (ORFs) 35–36, 55, 138
- operational genome annotation 135
- operons 138–139
- OptGene 310, 446
- optimal ATP production from other substrates 320
- optimal metabolic states, ATP production 360
- optimality principles 12
- optimization 298, 311
 - algorithms 447
 - biosynthetic networks 313–315
 - COBRA methods 298–299
 - extreme points 303, 340
 - finding functional states 300–301
 - genome-scale models 306–311
 - location of optimal solutions 303–304
 - metabolic network identification 309
 - network performance 312–315
 - and objective functions 298–300
 - sensitivity analysis 305–306
 - see also* alternative optimal solutions; equivalent optimal solutions; linear programming
- OptKnock 310, 444–446
- OptReg 310
- OptStrain 310
- ORFs (open reading frames) 35–36, 55, 138
- organelles 96
- orphan-filling genes, prediction methods 411, 413
- orphan reactions 409, 410–411
- orthonormal bases (SVD) 167, 186, 202–203, 220
 - bi-linear association reaction 193–196
 - concentration/flux combinations 189
 - fundamental subspaces of **S** 187
 - left null space 196
 - mathematical format 184
- output signatures 229
- oxygen effects 361, 362–364
- P/O ratio (phosphate/oxygen ratio) 42, 315, 366
- pan metabolic capabilities 59
- paradigm shifts 3
- paradigm, systems biology 8–11, 13–14, 453–455

- parallel processing 479
- parameter-free models 470–471
- parameters
 - diversity 255
 - dominant constraints 281
 - Escherichia coli* 68, 362–364, 379–383, 391–395
 - kinetics 296, 474
 - number of 281, 470
 - redundant constraints 281
 - restricted values 278, 279
 - thermodynamics 250
 - values 5, 277
 - varying two 368–371
 - see also* environmental parameters; genetic parameters; sensitivity analysis
- parasitism 87
- part lists 13–14, 455–458
- participation 179–180, 182
- participation numbers 159, 173, 190
- pathogenicity islands 430
- pathogens 80, 81
 - ALE studies 435–436
 - antibiotic effects 80
 - metabolism 78–80
 - reconstructions 80, 81
 - Salmonella* 429–431
 - Salmonella typhimurium* LT2 98, 99
 - simulations 108
- PathoLogic pathway hole filler 416–417
- pathway length 222–224, 225
- pathway length matrix 222
- pathway matrix (P) 215, 221–224
- pathway vectors 232
 - applications 221
 - crosstalk 229, 230
 - input–output relationships 229
 - reaction participation matrix 224–228
 - regulation 230, 231
- pathways 4, 18–19, 25, 204–205, 212, 220
 - analysis 250
 - closed paths 288
 - design 440
 - E. coli* 219, 312
 - predictor algorithms 414
 - reactions 212
 - red blood cells 217–219
 - redundancy 222
 - usage, sensitivity analysis 363, 364
 - see also* extreme pathways; glycolysis pathway; pathways as basis vectors; pentose phosphate pathway
- pathways as basis vectors 212–219
 - biochemical vs. mathematical dimensions 220
 - mathematics vs. biology 213
 - network-based pathway definitions 213
 - silent phenotypes 213
 - see also* extreme pathways
- PCA (Principal Component Analysis) 240
- PCC6803 *Synchocystis* strain 80–82
- pentose phosphate pathway (PPP) 22
 - combined stoichiometric matrix 22
 - coupling to glycolysis 22
 - non-oxidative branch 22
 - oxidative branch 22
 - robustness analysis 366
- perturbation experiments 463
- PFK (phosphofructokinase) enzyme 25–29
 - bound states 26
 - S matrix 27
- phenomics 146
- phenotypes
 - computation 31–32
 - generalist 259
 - see also* genotype–phenotype relationships
- phenotypic behaviors 270, 474, 476
- phenotypic functions 270, 468
- phenotypic instability 363, 364, 370
- phenotypic measures 427
- phenotypic phase planes (PhPP) 309, 369, 368–371, 377
 - genome-scale applications 371–377
 - objective functions 368, 370
 - phase characteristics 370–371
 - small-scale applications 371
- phenotypic phases
 - ATP production 360
 - finite number 368
 - futile phases 370
 - shadow prices 362
- phenotypic predictions 467
- phenotypic screens 384
- phenotypic states, cellular objectives 354
- PHFiller-GC 411
- phosphofructokinase (PFK) enzyme 25–29
- photon usage, *C. reinhardtii* 102–103
- photosynthesis
 - C. reinhardtii* 101
 - model organism 80–82
 - studying using GEMs 82
- PhPP *see* phenotypic phase planes
- phylogenetic coverage, prokaryotes 75–77
- phylogenetic profiles 417
- physical attributes of S 151

526 INDEX

- physical laws
 - absolute characteristics 12
 - Fick's law 251, 473
 - GEMs 472–473
 - Kirchoff's circuit laws 286, 288
 - mass conservation 12
 - Michaelis–Menten rate law 287
- physico-chemical causation 251
 - vs. biological 7, 254, 479
- physico-chemical constraints 12, 260–261, 285–289
- physico-chemical sciences
 - model-building approaches 255, 271
 - objective functions 344
 - perception of solution 470
- physics
 - causation 251
 - genetic code prediction 275
 - model-building approaches 255, 271–272
- pipelines 473
 - see also* workflows
- planning *see also* resources; workflows
- plasmids, naming conventions 55
- polyomic data 8
 - analysis 398, 449
 - constraints 296
 - genome metastructures 134–138, 144–145
 - integration 144–145
 - macrophage activation 405–406
 - pairwise integration 144, 146
 - time-dependent 474
 - two-species community 375–377
- polytope 208–209
- pools of variables 474, 480
- potential ORFs (pORFs) 138
- power-law distribution 175–178
- PPP *see* pentose phosphate pathway
- predictions
 - failures 407–409, 431
 - false negatives 69–70, 384, 409, 412–414, 416
 - false positives 70, 108, 384, 409, 415–416
 - true negatives 70, 384
 - true positives 70, 114, 384
- predictive methods 411–416, 420
 - see also* gap-filling; GEMs
- predictive practice 468
- Principal Component Analysis (PCA) 240
- Project K (*E. coli*) 73
- Project to Annotate 1000 Genomes 417
- prokaryotes 45, 75, 95
 - G. metallireducens* 91, 92, 375–377
 - G. sulfurreducens* 88, 94
 - genome-scale reconstructions 77
 - microbial communities 82–88
 - pathogens 78–80
 - phylogenetic coverage 75–77
 - Synechocystis* 80–82
- promiscuous enzymes 259, 406
- proteins
 - amino acids 275
 - binding sites 15, 118, 123
 - expression 296
 - folds 119–120
 - H. sapiens* 282
 - integrating with metabolic network reconstructions 130
 - interactions 146
 - Lrp regulon reconstruction 139–141
 - membrane capacity constraints 282
 - and metabolic models 293–295
 - missing information 407
 - multi-functionality 282
 - PFK enzyme 25–29
 - Protein Data Bank (PDB) 123
 - regulatory protein activity array 291
 - structure reconstructions 117–122, 139–141
 - synthesis 123, 130, 471, 474
 - T. maritima* 119–120
 - and their many states 25–29
 - thermal stability 120, 123
 - see also* gene–protein–reactions
- proteomics 15, 146
- protons 317, 320, 321
- proximal causation 11, 342–343, 354, 461
 - vs. distal causation 343
 - quantitative models 255
 - selection pressures 260
- Pseudomonas aeruginosa* 435
- Pseudomonas putida* 43
- pyruvate oxidoreductase (POR) 93
- QP (quadratic programming) 298, 299
- quality procedures 8, 11, 43, 306–308, 458
 - see also* curation; validation
- quantitative models 255–260
 - adjustable constraints 256
 - biological network properties 259–260
 - constraining behaviors 256, 257
 - constraints 256
 - distal causation 255–256
 - generations of models 257–259, 258
 - genome-scale 256–260
 - hierarchical nature of living systems 254
 - life sciences 255–256

- non-adjustable constraints 256
- physical sciences 255, 271
- proximal causation 255
- selection process 255–256
- solution spaces 256
- theory-based modeling 255
- see also* GEMs; *in silico* models; M-models; ME-Models
- quasi-steady states, time scale hierarchy and fast subsystems: quasi-steady states 287
- query tools 8, 249
- randomized sampling 233–234, 247, 300
 - Alzheimer's disease 244–247
 - E. coli* 237, 238
 - enzymopathies in red blood cells 241–243
 - exo-metabolome analysis 238–239
 - geometric objects 234
 - H. sapiens* 241–247
 - high-dimensional spaces 237–241
 - intracellular pathogens 243–244
 - low-dimensional spaces 234–237
 - M. tuberculosis* 243–244, 245
 - metabolic subtasking 244–247
 - mitochondria in the heart 241
 - Monte Carlo sampling 234, 242, 250, 301
 - parallelepipeds 235–237
 - redundant constraints 236
 - regulation modes 239, 240
 - rotating basis vectors 240–241
 - simple flux split 233–234
 - solution space volume 237
 - uniform random sampling 236
- ranking fitness, ALE endpoints 426–427
- RAVEN (reconstruction tool) 98
- reaction addition design 250
- reaction (link) adjacency matrix 179, 181
- reaction maps of **S** 161, 162, 163
- reaction pairs, partially coupled 227
- reaction participation 222
 - and connectivity 173–179
 - and correlated subsets 225–227, 228
 - matrix 224–228
 - participation number 173
- reactions 268
 - bi-linearity 193–196, 267–268
 - vs. compounds 163, 300
 - discovering new 462
 - glycolysis pathway 18–19
 - hot spots 121
 - linear networks 205–207
 - network gaps 409
 - non-negative linear basis 207–208
 - orphan 409
 - pathways 212
 - perturbation design 250
 - rates 285
 - and their stoichiometry 17–20
 - vectors 168, 199
 - see also* stoichiometric matrix, **S**; stoichiometry of chemical reactions
- Recon 1, for *Homo sapiens*
 - applications 106–108
 - functional testing of 325
 - high-throughput data 106–107
 - Homo sapiens*' metabolism 104–108
 - human kidney 123
 - imprinted genes 395
 - NADH 393
 - other mammals 114
- Recon X, for *Homo sapiens* 116
- reconstructed networks, automated tools 47
- reconstructions 8, 117, 133, 307
 - functional state combinatorics 268–270
 - integrating 128–130
 - jamborees 34, 98–99
 - protein structures 117–122
 - signaling networks 11, 12, 131–132, 133, 457
 - and statistical network models 398, 399
 - subsystems 417
 - teaching about 455–458
 - Tr/Tr networks 123–127
 - transcriptional regulatory networks (TRNs) 12
 - see also* BiGG information; bottom-up reconstruction; bottom-up/top-down reconstruction; *Escherichia coli*; gap-filling; genome-scale reconstructions; metabolic network reconstructions; Recon 1; Recon X; workflows
- red blood cell metabolism
 - enzymopathies 241–243
 - extreme pathways 338
 - pathways 218, 217–219
- redox potential, production of 320–325
- reduced costs 305, 317–319
- redundancy 480
 - AOS 338–339
 - COBRA 470
 - equivalent reaction sets 335
 - FVA 335
 - gene removal 379, 380
 - genome-scale 259
 - Geobacter* spp. 92, 93

528 INDEX

- mathematical definitions 480
- pathways 222, 338
- redundant constraints 281–282, 454
 - low-dimensional spaces 236
 - steady-state flux vector, confining 210–211
- regulatory constraints 289–291
 - biological perspective 262
 - consequences of 222
 - missing 474
 - picking candidate states 273
 - as restraints 212, 230
 - transcriptional 290–291
- regulatory elements 139
- regulatory methods 250
- regulatory off/on modification (ROOM) 309, 388–390
- regulatory protein activity array 291
- regulon reconstruction 139, 140
- relative reaction rates 267
- relative thinking 12
- reporter metabolites 402
- resources, network reconstruction process 43–45
- restraint, regulation as 212, 230, 231
- restricted parameter values, quantifying 278, 279
- reversible bi-molecular reaction 180
- reversible conversion 154
- reversible reaction 180
- rFBA 309
- Rhodoferrax ferrireducens* 86
- right null space of **S** *see* null space of **S**
- right singular vectors 185, 187
- RNA polymerase binding regions (RBRs) 138
- RNAP-guided transcript segments (RTs) 138
- RobustKnock 446
- robustness analysis 309
 - applications 366–368
 - enzymes 366
 - mathematical definitions 480
 - oxygen uptake rate, effects on growth rate 364
 - sensitivity analysis 364–366
 - varying two parameters 368–371
- ROOM (regulatory off/on modification) 309, 388–390
- root no-consumption metabolites 410
- root no-production metabolites 410
- row space of **S** 167, 168
- S** matrix *see* stoichiometric matrix
- Saccharomyces cerevisiae* (baker's yeast) 96
 - applications 98
 - community-based reconstruction 98–101
 - consensus reconstructions 100–101
 - flux coupling finder (FCF) 228
 - GEM history 96
 - gene–gene interactions 393
 - gene knock-out 383–384
 - genome 45, 64, 97
 - metabolite coupling 182
 - objective functions 353
 - PhPP 374
 - reconstruction 96–98
 - Salmonella* 429–431
 - typhimurium* LT2 98, 99
 - scale-free networks 175
 - scope gaps 410
 - selection pressures
 - ALE studies 423, 428
 - distal causation 260, 342
 - dominant constraints 282
 - GEM axioms 12
 - natural selection 273
 - objective functions 12, 342, 355
 - PhPP 370
 - proximal causation 260
 - quantitative models 255–256
 - synthetic biology 431
 - self 260, 472, 480
 - self-imposed constraints 262
 - sensitivity analysis 305–306, 364, 366
 - shadow prices 305, 317–319, 362, 363
 - sigma factor network 141
 - signal transduction 260
 - signaling molecules 282–284
 - signaling networks 11, 12, 131–132, 133, 457
 - silent phenotypes 213, 259, 329
 - simple flux splits 211, 233–234
 - simultaneity 6, 296, 325, 338
 - single gene knock-outs 378–379
 - algorithmic formalization 385
 - classifying outcomes 384
 - E. coli* 31, 31–32, 379–383
 - genetic parameters 378–379
 - genome-scale studies 383–386
 - H. influenzae* 383
 - large-scale deletion analysis 383–384
 - lethality 380–383
 - mouse studies 386
 - non-lethal 386–390
 - S. cerevisiae* 383–384
 - strain growth properties 383–386
 - single nucleotide polymorphisms *see* SNPs
 - singular value decomposition (SVD) 184–188
 - bi-linear association reaction 193–196
 - biomass formation 197–199

- chemical reactions analogy 190
- concentration/flux combinations 189
- content of U , Σ , and V 185–188
- elementary reactions 191–196
- genome-scale matrices 199
- key properties 188
- linear reversible reactions 191–193
- mapping between singular vectors 188
- mathematical format 184
- matrices 184–185
- metabolic network 196–199
- mode-by-mode reconstruction of S 188
- network structure 196–199
- nomenclature 185
- reaction network properties 189–191
- of S 184–185, 191, 194–195
- as a series of transformations 185
- singular value spectrum 185–187
- systems reactions 189, 190
- see also* orthonormal bases
- SMILEY 411, 412–414
- SNPs (single nucleotide polymorphisms),
mapping causal SNPs 396
- solution spaces, confining
 - computing size and contents 234
 - constraints and 296
 - geometric objects 234
 - high-throughput data 401
 - restraints, regulatory 230
 - see also* constraints; gene knock-outs; null
space of S ; randomized sampling;
steady-state flux vector, confining
- sparse matrices 172
- spatial constraints 261
- species level
 - DIET 375–377
 - E. coli* reconstruction 59–64
 - interspecies interaction 474, 476
 - multi-species relationships 77
 - phenotypic characteristics 152
 - two-species community 375–377
- specific/abstract knowledge 268
- specificity, enzymes 117–119, 406
- SR-FBA 310
- stable isotope tracers 309
- standard operating procedure (SOP), network
reconstruction process 43–45
- Staphylococcus aureus* 182
- state variables 399
- statistical models 268, 398, 399
- steady-state assumptions
 - COBRA use 469
 - constraints from chemistry 12, 286
 - multi-scale analysis 472
- steady-state flux distributions, S matrix 167
- steady-state flux vector, confining 208–212
 - closed spaces 208–209
 - convex vs. linear bases 208
 - dominant constraints 210–211
 - importance of constraints 210–212
 - polytope 208–209
 - redundant constraints 210–211
 - varying constraints 211–212
- steady vs. dynamic states 469
- stimulons, reconstructing 140–141
- stoichiometric coefficients 151, 171, 190
- stoichiometric matrix, S 12–13, 149, 151, 171
 - attributes 151–153
 - basic chemistry 155–158
 - BiGG information 40–42
 - binary form 172–173
 - bound states of PFK 27
 - combining 155
 - compound maps 161, 162–164
 - as a connectivity matrix 153, 161–165
 - as a data matrix 153–161
 - distal causation 152
 - dynamic mass balances 168, 169
 - elementary reactions 154–155
 - fundamental subspaces 166–167
 - genome annotation 13–14
 - glycolysis 158–161
 - as a linear transformation 165–168
 - mapping fluxes onto concentration time
derivatives 165–166
 - mathematical attributes 152
 - and network models 168, 170
 - numerical attributes of S 153
 - reaction maps 161, 162, 163
 - rearranging 174–175
 - sparse matrix, S as 153
 - systems boundary 169–171
 - systemic attributes 153
 - systems science 168, 170
- stoichiometry of chemical reactions 17–20, 149,
266
 - constraints 314
 - sensitivity analysis 363, 364, 366
- strain design 250, 300, 445–446, 448
 - see also* metabolic engineering
- structural biology 119–122
- structural genome annotation 135
- Subliminal Toolbox 47
- subspaces of S *see* fundamental subspaces of S

530 INDEX

- subsystems approach, to gene annotation 417–418
- survival 270
- SVD *see* singular value decomposition
- Synechocystis*
 - iJN678 reconstruction 82
 - PCC6803 strain 78, 80–82
- synthetic biology 48, 431
 - see also* metabolic engineering
- synthetic gene interactions 392–393
- synthetic lethality 378, 390–395
- system 19
- system boundaries
 - GEM formulation 68, 308, 457
 - reconstructed networks 19, 169–170, 471
 - S** matrix 169–171
 - scope gaps 410
 - Tr/Tr networks 127
 - virtual 169, 215
- system decomposition 472
- system perturbation variables 428
- systems biology 1, 3,
 - ALE outcomes 427
 - applications 454
 - bottom-up rationale 2, 275, 453
 - causation 7, 253–255
 - core paradigm 8–11, 13–14, 453–455
 - CoSy 85
 - foundations 73
 - genome-scale science 478–479
 - multidisciplinary approach 464
 - mutations 252, 432
 - questions in 471–472
 - terminology 464
- systems biology, teaching 453, 465, 466
 - building a curriculum 464
 - computing functional states of networks 458–462
 - core paradigm 453–455
 - high-throughput technologies 455
 - network reconstruction 455–458
 - prospective experimentation 462–464
- systems science 11
- systems viewpoints, vs. components 3, 264–265
- tangles of cycles 267
- target identification 393
- TCA cycle (tricarboxylic acid cycle) 366, 393
- teaching systems biology *see* systems biology, teaching
- template reactions 126
- temporal–spatial organization, biological networks 271
- theory-based modeling 255, 271–272
- thermal noise 479
- thermal stability 120–123
- thermodynamics 250, 287–288
 - constraints 296
 - driving forces 167, 201
 - metabolic flux analysis 309
 - S** matrix 167
- thermophiles 130
- Thermotoga maritima*
 - genome-scale integration 120
 - integrating reconstructions 130
 - protein folds 119–120
 - subsystems 417–418
- threonine strain 367–368
- time-dependency
 - ALE studies 431
 - biological systems 255, 265
 - objective functions 431
 - polyomic data sets 474
 - third-generation GEMs 257–259, 474
- time derivatives, **S** matrix 167
- time-invariants, **S** matrix 167
- time variance 262, 270
- timescales, manual curation 43
- tissue type models 111–114
- top-down/bottom-up reconstruction 138, 275, 454
 - see also* bottom-up reconstruction
- topological constraints (molecular) 11
- topology, network *see* network topology
- Tr/Tr (transcription and translation) networks 123–127
 - automated reconstruction 126
 - cell growth 127
 - co-sets 127
 - curation 126
 - network modules 127
 - omics data mapping 127
 - omics data types 127
 - reconstruction applications 127
 - reconstruction workflow 125–127
 - topological properties 127
- transcription binding 474
- transcription factors (TFs) 139, 404–405
- transcription units (TUs) 134, 138
- transcriptional regulators, identifying 130
- transcriptional regulatory networks (TRNs) 11, 134, 138–142, 290–291, 474
- transcriptomics 12, 15, 131, 146

- transferred properties 268
- tricarboxylic acid cycle (TCA cycle) 366, 393
- TRNs (transcriptional regulatory networks) 11, 134, 138–142, 290–291, 474
- true-negative predictions 70, 384
- true-positive predictions 70, 114, 384
- tuning expression 446

- unbiased biological network state assessments 221
- unbiased COBRA methods 250, 299–300, 470
- unbiased network modules 127
- unbiased network reconstructions 456, 457
- unbiased searches, for objective functions 353
- unbounded solutions 305
- uniform random sampling 233–234, 309
- unique solutions, linear programming problems 304

- validation (network reconstruction) 8, 41, 42, 457
- variable pools 474, 480
- vector spaces *see* basis for vector spaces; fundamental subspaces of **S**; pathway vectors; pathways as basis vectors; steady-state flux vector, confining
- Vibrio cholerae* 417
- virtual boundaries 169, 215
- virulence 404–405
- visualization tools 301

- ‘who-listens-to-whom’ 260, 479
- wiring diagrams 13–14
- workflows
 - automating 43–45
 - basic reconstruction 35–42, 49, 457
 - cell factory design 443–446, 445
 - complex relationships 11
 - CoSy biology 85
 - future work 133
 - G. sulfurreducens* 90
 - Geobacter* spp. 91
 - iterations 44, 49, 51–65, 407–409, 474, 475
 - multi-cell reconstructions 108–111
 - polyomic data integration 145
 - reconstruction jamborees 98–99
 - refactoring DNA for synthetic biology 142, 144
 - regulons 140
 - Salmonella* 99
 - sampling allowable states 241
 - sigma factor network 141
 - SOPs 11, 43, 45
 - synthetic biology, refactoring DNA 142–144
 - systems biology 8, 474, 475
 - systems metabolic engineering 440, 443
 - Tr/Tr networks 125–126
 - see also* conversion processes; gap-filling
- yeast *see* *Saccharomyces cerevisiae*