

References

- Akashi H, Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 99:3695–3700
- Allen A, Flemstrom G, Garner A, Kivilaakso E (1993) Gastroduodenal mucosal protection. *Physiol Rev* 73:823–857
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Argos P, Rossmann MG, Grau UM, Zuber A, Franck G, Tratschin JD (1979) Thermal stability and protein structure. *Biochemistry (Mosc)* 18:5698–5703
- Aris-Brosou S, Xia X (2008) Phylogenetic analyses: a toolbox expanding towards Bayesian methods. *Int J Plant Genomics* 2008:683509
- Bader DA, Moret BM, Yan M (2001) A linear-time algorithm for computing inversion distance between signed permutations with an experimental study. *J Comput Biol* 8:483–491
- Bahir I, Fromer M, Prat Y, Linial M (2009) Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Mol Syst Biol* 5:311
- Baik SC, Kim KM, Song SM, Kim DS, Jun JS, Lee SG, Song JY, Park JU, Kang HL, Lee WK, Cho MJ, Youn HS, Ko GH, Rhee KH (2004) Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Helicobacter pylori* strain 26695. *J Bacteriol* 186:949–955
- Ballester R, Marchuk D, Boguski M, Saulino A, Letcher R, Wigler M, Collins F (1990) The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. *Cell* 63:851–859
- Barker D, Meade A, Pagel M (2007) Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. *Bioinformatics* 23:14–20
- Barker D, Pagel M (2005) Predicting functional gene links from phylogenetic-statistical analyses of whole genomes. *PLoS Comput Biol* 1:e3
- Bauerfeind P, Garner R, Dunn BE, Mobley HL (1997) Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH. *Gut* 40:25–30
- Baumgartner HK, Montrose MH (2004) Regulated alkali secretion acts in tandem with unstirred layers to regulate mouse gastric surface pH. *Gastroenterology* 126:774–783
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 57:289–300
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple hypothesis testing under dependency. *Ann Stat* 29:1165–1188
- Berman P, Hannenhalli S, Karpinski M (2002) 1.375-approximation algorithm for sorting by reversals. *Eur Symp Algorithms* 2461:401–408

- Bestor TH, Coxon A (1993) The pros and cons of DNA methylation. *Curr Biol* 6:384–386
- Bogenhagen DF, Clayton DA (2003) The mitochondrial DNA replication bubble has not burst. *Trends Biochem Sci* 28:357–360
- Brown CJ, Garner EC, Keith Dunker A, Joyce P (2001) The power to detect recombination using the coalescent. *Mol Biol Evol* 18:1421–1424
- Brown TA, Cecconi C, Tkachuk AN, Bustamante C, Clayton DA (2005) Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev* 19:2466–2476
- Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172:2665–2681
- Bulmer M (1987) Coevolution of codon usage and transfer RNA abundance. *Nature* 325:728–730
- Bulmer M (1991) The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129:897–907
- Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268:78–94
- Burge CB, Karlin S (1998) Finding the genes in genomic DNA. *Curr Opin Struct Biol* 8:346–354
- Bury-Mone S, Skouloubri S, Labigne A, De Reuse H (2001) The *Helicobacter pylori* UreI protein: role in adaptation to acidity and identification of residues essential for its activity and for acid activation. *Mol Microbiol* 42:1021–1034
- Cardon LR, Burge C, Clayton DA, Karlin S (1994) Pervasive CpG suppression in animal mitochondrial genomes. *Proc Natl Acad Sci U.S.A* 91:3799–3803
- Carullo M, Xia X (2008) An extensive study of mutation and selection on the wobble nucleotide in tRNA anticodons in fungal mitochondrial genomes. *J Mol Evol* 66:484–493
- Chambaud I, Heilig R, Ferris S, Barbe V, Samson D, Galisson F, Moszer I, Dybvig K, Wroblewski H, Viari A, Rocha EP, Blanchard A (2001) The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. *Nucleic Acids Res* 29:2145–2153
- Chan PP, Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucl Acids Res* 37:D93–97
- Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* 28:693–705
- Clayton DA (2000) Transcription and replication of mitochondrial DNA. *Hum Reprod* 15:11–17
- Colby C, Edlin G (1970) Nucleotide pool levels in growing, inhibited, and transformed chick fibroblast cells. *Biochemistry (Mosc)* 9:917
- Crick FH (1966) Codon–anticodon pairing: the wobble hypothesis. *J Mol Biol* 19:548–555
- Curran JF (1995) Decoding with the A:I wobble pair is inefficient. *Nucleic Acids Res* 23:683–688
- Dalgaard JZ, Garrett RA (1993) Archaeal hyperthermophile genes. In: Kates M, Kushner DJ, Matheson AT (eds) *The biochemistry of Archaea (Archaeobacteria)*. Elsevier, Amsterdam
- Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
- Engel E, Peskoff A, Kauffman GL Jr, Grossman MI (1984) Analysis of hydrogen ion concentration in the gastric gel mucus layer. *Am J Physiol* 247:G321–G338
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368–376
- Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15
- Felsenstein J (2002) PHYLIP 3.6 (phylogeny inference package). Version 3.6. Department of Genetics, University of Washington, Seattle
- Felsenstein J (2004) *Inferring phylogenies*. Sinauer, Sunderland
- Frederico LA, Kunkel TA, Shaw BR (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry (Mosc)* 29:2532–2537
- Frederico LA, Kunkel TA, Shaw BR (1993) Cytosine deamination in mismatched base pairs. *Biochemistry (Mosc)* 32:6523–6530
- Galtier N, Lobry JR (1997) Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* 44:632–636

- Gardy JL, Spencer C, Wang K, Ester M, Tuszynski GE, Simon I, Hua S, deFays K, Lambert C, Nakai K, Brinkman FS (2003) PSORT-B: Improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res* 31:3613–3617
- Ge Y, Sealfon SC, Speed TP (2008) Some step-down procedures controlling the false discovery rate under dependence. *Stat Sin* 18:881–904
- Gojbori T, Li WH, Graur D (1982) Patterns of nucleotide substitution in pseudogenes and functional genes. *J Mol Evol* 18:360–369
- Goto M, Washio T, Tomita M (2000) Causal analysis of CpG suppression in the *Mycoplasma* genome. *Microb Comp Genomics* 5:51–58
- Gouy M, Gautier C (1982) Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Res* 10:7055–7064
- Grosjean H, de Crecy-Lagard V, Marck C (2010) Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett* 584:252–264
- Haas J, Park E-C, Seed B (1996) Codon usage limitation in the expression of HIV-1 envelope glycoprotein. *Curr Biol* 6:315–324
- Hamajima N, Goto Y, Nishio K, Tanaka D, Kawai S, Sakakibara H, Kondo T (2004) *Helicobacter pylori* eradication as a preventive tool against gastric cancer. *Asian Pac J Cancer Prev* 5:246–252
- Hanada K, Suzuki Y, Gojbori T (2004) A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. *Mol Biol Evol* 21:1074–1080
- Harvey PH, Pagel MD (1991) *The comparative method in evolutionary biology*. Oxford University Press, Oxford
- Hey J (2000) Human mitochondrial DNA recombination: can it be true? *Trends Ecol Evol* 15:181–182
- Higgs PG, Ran W (2008) Coevolution of codon usage and tRNA genes leads to alternative stable states of biased codon usage. *Mol Biol Evol* 25:2279–2291
- Hofer A, Steverding D, Chabes A, Brun R, Thelander L (2001) *Trypanosoma brucei* CTP synthetase: a target for the treatment of African sleeping sickness. *Proc Natl Acad Sci U S A* 98:6412–6416
- Hughes AL (2000) Polyploidization and vertebrate origins: a review of the evidence. In: Sankoff D, Nadeau JH (eds) *Comparative genomics*. Kluwer Academic Publishers, Netherland, pp 493–502
- Hunt RH (2004) Will eradication of *Helicobacter pylori* infection influence the risk of gastric cancer? *Am J Med* 117:86S–91S
- Hurst LD, Merchant AR (2001) High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes. *Proc R Soc Lond B* 268:493–497
- Husmeier D, Wright F (2005) Detecting recombination in DNA sequence alignments. In: Husmeier D, Dybowski R, Roberts S (eds) *Probabilistic modeling in bioinformatics and medical informatics*. Springer, London, p 504
- Ikemura T (1981a) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J Mol Biol* 146:1–21
- Ikemura T (1981b) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E coli* translational system. *J Mol Biol* 151:389–409
- Ikemura T (1982) Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. Differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. *J Mol Biol* 158:573–597
- Ikemura T (1992) Correlation between codon usage and tRNA content in microorganisms. In: Hatfield DL, Lee BJ, Pirtle RM (eds) *Transfer RNA in protein synthesis*. CRC Press, Boca Raton, pp 87–111
- Irimia M, Penny D, Roy SW (2007) Coevolution of genomic intron number and splice sites. *Trends Genet* 23:321

- Jacob F (1988) *The statue within: an autobiography*. Basic Books, Inc, New York
- Jakobsen IB, Easteal S (1996) A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. *Comput Appl Biosci* 12:291–295
- Jenkins GM, Holmes EC (2003) The extent of codon usage bias in human RNA viruses and its evolutionary origin. *Virus Res* 92:1–7
- Jia W, Higgs PG (2008) Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. *Mol Biol Evol* 25:339–351
- Josse J, Kaiser AD, Kornberg A (1961) Enzymatic synthesis of deoxyribonucleic acid VII. Frequencies of nearest neighbor base-sequences in deoxyribonucleic acid. *J Biol Chem* 236:864–875
- Karlin S, Burge C (1995) Dinucleotide relative abundance extremes: a genomic signature. *TIG* 11:283–290
- Karlin S, Mrazek J (1996) What drives codon choices in human genes. *J Mol Biol* 262:459–472
- Keating CP, Hill MK, Hawkes DJ, Smyth RP, Isel C, Le SY, Palmenberg AC, Marshall JA, Marquet R, Nabel GJ, Mak J (2009) The A-rich RNA sequences of HIV-1 pol are important for the synthesis of viral cDNA. *Nucleic Acids Res* 37:945–956
- Kececioglu JD, Sankoff D (eds) (1994) In: *In Proceedings of the 5th symposium on combinatorial pattern matching*
- Kececioglu JD, Sankoff D (1995) Exact and approximation algorithms for sorting by reversal. *Algorithmica* 13:180–210
- Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49:725–738
- Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in *Hominoidea*. *J Mol Evol* 29:170–179
- Kliman RM, Bernal CA (2005) Unusual usage of AGG and TTG codons in humans and their viruses. *Gene* 352:92
- Kraytsberg Y, Schwartz M, Brown TA, Ebralidse K, Kunz WS, Clayton DA, Vissing J, Khrapko K (2004) Recombination of human mitochondrial DNA. *Science* 304:981
- Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9:299–306
- Kushiro A, Shimizu M, Tomita K-I (1987) Molecular cloning and sequence determination of the *tuf* gene coding for the elongation factor Tu of *Thermus thermophilus* HB8. *Eur J Biochem* 170:93–98
- Kypr J, Mrazek JAN (1987) Unusual codon usage of HIV. *Nature* 327:20
- Lemey P, Posada D (2009) Introduction to recombination detection. In: Lemey P, Salemi M, Vandamme AM (eds) *The phylogenetic handbook*. Cambridge University Press, Cambridge, pp 493–518
- Li W-H (1983) *Evolution of duplicate genes and pseudogenes*. Sinauer, Sunderland
- Li WH, Gojobori T, Nei M (1981) Pseudogenes as a paradigm of neutral evolution. *Nature* 292:237–239
- Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362:709–715
- Lobry JR (1996) Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol* 13:660–665
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 73:152–160
- Lopez P, Philippe H, Myllykallio H, Forterre P (1999) Identification of putative chromosomal origins of replication in *Archaea*. *Mol Microbiol* 32:883–886
- Lough J, Jackson M, Morris R, Moyer R (2001) Bisulfite-induced cytosine deamination rates in *E. coli* SSB:DNA complexes. *Mutat Res* 478:191–197
- Ma P, Xia X (2011) Factors affecting splicing strength of yeast genes. *Comp Funct Genomics* Article ID 212146:13

- Marin A, Xia X (2008) GC skew in protein-coding genes between the leading and lagging strands in bacterial genomes: new substitution models incorporating strand bias. *J Theor Biol* 253:508–513
- Marín A, Xia X (2008) GC skew in protein-coding genes between the leading and lagging strands in bacterial genomes: new substitution models incorporating strand bias. *J Theor Biol* 253:508–513
- Martinez MA, Vartanian J-P, Simon W-H (1994) Hypermutagenesis of RNA Using human immunodeficiency virus type 1 reverse transcriptase and biased dNTP concentrations. *Proc Natl Acad Sci U S A* 91:11787–11791
- Martins EP, Hansen TF (1997) Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am Nat* 149:646–667
- Matin A, Zychlinsky E, Keyhan M, Sachs G (1996) Capacity of *Helicobacter pylori* to generate ionic gradients at low pH is similar to that of bacteria which grow under strongly acidic conditions. *Infect Immun* 64:1434–1436
- Menaker RJ, Sharaf AA, Jones NL (2004) *Helicobacter pylori* infection and gastric cancer: host, bug, environment, or all three? *Curr Gastroenterol Rep* 6:429–435
- Menz GL, Hazell SL (1996) The urea cycle of *Helicobacter pylori*. *Microbiology* 142:2959–2967
- Menz GL, Jimenez BM, Hazell SL, Gero AM, O'Sullivan WJ (1994) De novo synthesis of pyrimidine nucleotides by *Helicobacter pylori*. *J Appl Bacteriol* 77:1–8
- Mine T, Muraoka H, Saika T, Kobayashi I (2005) Characteristics of a clinical isolate of urease-negative *Helicobacter pylori* and its ability to induce gastric ulcers in *Mongolian gerbils*. *Helicobacter* 10:125–131
- Mobley HL, Hu LT, Foxal PA (1991) *Helicobacter pylori* urease: properties and role in pathogenesis. *Scand J Gastroenterol Suppl* 187:39–46
- Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci U S A* 93:10268–10273
- Muto A, Osawa S (1987) The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc Natl Acad Sci USA* 84:166–169
- Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 24:34–36
- Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res* 28:292
- Nakashima H, Fukuchi S, Nishikawa K (2003) Compositional changes in RNA, DNA and proteins for bacterial adaptation to higher and lower temperatures. *J Biochem (Tokyo)* 133:507–513
- Ngumbela KC, Ryan KP, Sivamurthy R, Brockman MA, Gandhi RT, Bhardwaj N, Kavanagh DG (2008) Quantitative effect of suboptimal codon usage on translational efficiency of mRNA encoding HIV-1 gag in Intact T Cells. *PLoS ONE* 3:e2356
- Nichols T, Hayasaka S (2003) Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat Meth Med Res* 12:419–446
- Nur I, Szyf M, Razin A, Glaser G, Rottem S, Razin S (1985) Procaryotic and eucaryotic traits of DNA methylation in *Spiroplasma (Mycoplasmas)*. *J Bacteriol* 164:19–24
- Nussinov R (1984) Doublet frequencies in evolutionary distinct groups. *Nucleic Acids Res* 12:1749–1763
- Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304
- Page RDM (2003) Introduction. In: Page RDM (ed) *Tangled trees: phylogeny, cospeciation and coevolution*. University of Chicago Press, Chicago, pp 1–21
- Pagel M (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc R Soc Lond B Biol Sci* 255:37–45
- Pagel M (1997) Inferring evolutionary processes from phylogenies. *Zool Scr* 26:331–348
- Pagel M (1999) Inferring the historical patterns of biological evolution. *Nature* 401:877–884

- Palidwor GA, Perkins TJ, Xia X (2010) A general model of codon bias due to GC mutational bias. *PLoS One* 5:e13431
- Paterson AM, Gray RD, Wallis GP (1995) Of lice and men: the return of the 'comparative parasitology' debate. *Parasitol Today* 11:158–160
- Plotkin JB, Kudla G (2010) Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet* 12:32–42
- Plotkin JB, Robins H, Levine AJ (2004) Tissue-specific codon usage and the expression of human genes. *Proc Natl Acad Sci U S A* 101:12588–12591
- Posada D (2002) Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Mol Biol Evol* 19:708–717
- Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci U S A* 98:13757–13762
- Press WH, Teukolsky SA, Tetterling WT, Flannery BP (1992) *Numerical recipes in C: the art of scientific computing*. Cambridge University Press, Cambridge
- Razin A, Razin S (1980) Methylated bases in mycoplasmal DNA. *Nucleic Acids Res* 8:1383–1390
- Rektorschek M, Buhmann A, Weeks D, Schwan D, Bensch KW, Eskandari S, Scott D, Sachs G, Melchers K (2000) Acid resistance of *Helicobacter pylori* depends on the UreI membrane protein and an inner membrane proton barrier. *Mol Microbiol* 36:141–152
- Rideout WMI, Coetzee GA, Olumi AF, Jones PA (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 249:1288–1290
- Rimsky L, Hauber J, Dukovich M, Malim MH, Langlois A, Cullen BR, Greene WC (1988) Functional replacement of the HIV-1 rev protein by the HTLV-1 rex protein. *Nature* 335:738–740
- Rocha EP, Danchin A (2002) Base composition bias might result from competition for metabolic resources. *Trends Genet* 18:291–294
- Sachs G, Meyer-Rosberg K, Scott DR, Melchers K (1996) Acid, protons and *Helicobacter pylori*. *Yale J Biol Med* 69:301–316
- Sachs G, Weeks DL, Melchers K, Scott DR (2003) The gastric biology of *Helicobacter pylori*. *Annu Rev Physiol* 65:349–369
- Saenger W (1984) *Principles of nucleic acid structure*. Springer, New York
- Salemi M, Gray RR, Goodenow MM (2008) An exploratory algorithm to identify intra-host recombinant viral sequences. *Mol Phylogenet Evol* 49:618
- Salminen M, Martin D (2009) Detecting and characterizing individual recombination events. In: Lemey P, Salemi M, Vandamme AM (eds) *The phylogenetic handbook*. Cambridge University Press, Cambridge, pp 519–548
- Salminen MO, Carr JK, Burke DS, McCutchan FE (1995) Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* 11:1423–1425
- Sancar A, Sancar GB (1988) DNA repair enzymes. *Annu Rev Biochem* 57:29–67
- Sansonetti PJ, d'Hauteville H, Formal SB, Toucas M (1982a) Plasmid-mediated invasiveness of "Shigella-like" *Escherichia coli*. *Ann Microbiol (Paris)* 133:351–355
- Sansonetti PJ, Kopecko DJ, Formal SB (1982b) Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun* 35:852–860
- Schluter D, Price TD, Mooers AØ, Ludwig D (1997) Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699–1711
- Scott D, Weeks D, Melchers K, Sachs G (1998) The life and death of *Helicobacter pylori*. *Gut* 43:S56–60
- Scott DR, Marcus EA, Weeks DL, Sachs G (2002) Mechanisms of acid resistance due to the urease system of *Helicobacter pylori*. *Gastroenterology* 123:187–195
- Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem* 66:409–435
- Sharp PM (1986) What can AIDS virus codon usage tell us? *Nature* 324:114
- Sharp PM, Li WH (1987) The codon adaptation Index a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15:1281–1295

- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116
- Siavoshi F, Malekzadeh R, Daneshmand M, Smoot DT, Ashktorab H (2004) Association between *Helicobacter pylori* infection in gastric cancer, ulcers and gastritis in Iranian patients. *Helicobacter* 9:470
- Singer CE, Ames BN (1970) Sunlight ultraviolet and bacterial DNA base ratios. *Science* 170:822–826
- Slonczewski JL, Rosen BP, Alger JR, Macnab RM (1981) pH homeostasis in *Escherichia coli*: measurement by ³¹P nuclear magnetic resonance of methylphosphonate and phosphate. *Proc Natl Acad Sci U S A* 78:6271–6275
- Stingl K, Altendorf K, Bakker EP (2002a) Acid survival of *Helicobacter pylori*: how does urease activity trigger cytoplasmic pH homeostasis? *Trends Microbiol* 10:70–74
- Stingl K, Uhlemann E-M, Schmid R, Altendorf K, Bakker EP (2002b) Energetics of *Helicobacter pylori* and its implications for the mechanism of urease-dependent acid tolerance at pH 1. *J Bacteriol* 184:3053–3060
- Stingl K, Uhlemann EM, Deckers-Hebestreit G, Schmid R, Bakker EP, Altendorf K (2001) Prolonged survival and cytoplasmic pH homeostasis of *Helicobacter pylori* at pH 1. *Infect Immun* 69:1178–1180
- Stoebel DM (2005) Lack of evidence for horizontal transfer of the lac operon into *Escherichia coli*. *Mol Biol Evol* 22:683–690
- Strebel K (2005) APOBEC3G & HTLV-1: inhibition without deamination. *Retrovirology* 2:37
- Suchard MA, Weiss RE, Dorman KS, Sinsheimer JS (2002) Oh brother, where art thou? A Bayes factor test for recombination with uncertain heritage. *Syst Biol* 51:715–728
- Sueoka N (1964) On the evolution of informational macromolecules. Academic Press, New York
- Sved J, Bird A (1990) The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc Natl Acad Sci USA* 87:4692–4696
- Swofford DL (2000) Phylogenetic analysis using parsimony (* and other methods). Sinauer, Sunderland
- Tanaka M, Ozawa T (1994) Strand asymmetry in human mitochondrial DNA mutations. *Genomics* 22:327–335
- Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Venter JC et al (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388:539–547
- Valenzuela M, Cerda O, Toledo H (2003) Overview on chemotaxis and acid resistance in *Helicobacter pylori*. *Biol Res* 36:429–436
- Van Dooren S, Pybus OG, Salemi M, Liu HF, Goubau P, Remondegui C, Talarmin A, Gotuzzo E, Alcantara LC, Galvao-Castro B, Vandamme AM (2004) The low evolutionary rate of human T-cell lymphotropic virus type-1 confirmed by analysis of vertical transmission chains. *Mol Biol Evol* 21:603–611
- van Hemert FJ, Berkhout B (1995) The tendency of lentiviral open reading frames to become A-rich: constraints imposed by viral genome organization and cellular tRNA availability. *J Mol Evol* 41:132–140
- van Weringh A, Ragonnet-Cronin M, Pranckeviciene E, Pavon-Eternod M, Kleiman L, Xia X (2011) HIV-1 modulates the tRNA pool to improve translation efficiency. *Mol Biol Evol* 28:1827–1834
- Vartanian J-P, Henry M, Wain-Hobson S (2002) Sustained G→A hypermutation during reverse transcription of an entire human immunodeficiency virus type 1 strain Vau group O genome. *J Gen Virol* 83:801–805
- Vinci G, Xia X, Veitia RA (2008) Preservation of genes involved in sterol metabolism in cholesterol auxotrophs: facts and hypotheses. *PLoS ONE* 3:e2883
- Vos RA, Balhoff JP, Caravas JA, Holder MT, Lapp H, Maddison WP, Midford PE, Priyam A, Sukumaran J, Xia X, Stoltzfus A (2012) NeXML: rich, extensible, and verifiable representation of comparative data and metadata. *Syst Biol* 61:675–689

- Wang HC, Hickey DA (2002) Evidence for strong selective constraint acting on the nucleotide composition of 16S ribosomal RNA genes. *Nucleic Acids Res* 30:2501–2507
- Wang HC, Xia X, Hickey DA (2006) Thermal adaptation of ribosomal RNA genes: a comparative study. *J Mol Evol* 63:120–126
- Weber MJ, Edlin G (1971) Phosphate transport, nucleotide pools, and ribonucleic acid synthesis in growing and in density-inhibited 3T3 cells. *J Biol Chem* 246:1828–1833
- Weeks DL, Eskandari S, Scott DR, Sachs G (2000) A H⁺-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization. *Science* 287:482–485
- Wen Y, Marcus EA, Matrubutham U, Gleeson MA, Scott DR, Sachs G (2003) Acid-adaptive genes of *Helicobacter pylori*. *Infect Immun* 71:5921–5939
- Williams CL, Preston T, Hossack M, Slater C, McColl KE (1996) *Helicobacter pylori* utilises urea for amino acid synthesis. *FEMS Immunol Med Microbiol* 13:87–94
- Wiuf C, Christensen T, Hein J (2001) A simulation study of the reliability of recombination detection methods. *Mol Biol Evol* 18:1929–1939
- Wright F (1990) The ‘effective number of codons’ used in a gene. *Gene* 87:23–29
- Xia X (1996) Maximizing transcription efficiency causes codon usage bias. *Genetics* 144:1309–1320
- Xia X (1998a) How optimized is the translational machinery in *Escherichia coli*, *Salmonella typhimurium* and *Saccharomyces cerevisiae*? *Genetics* 149:37–44
- Xia X (1998b) The rate heterogeneity of nonsynonymous substitutions in mammalian mitochondrial genes. *Mol Biol Evol* 15:336–344
- Xia X (2001) Data analysis in molecular biology and evolution. Kluwer Academic Publishers, Boston
- Xia X (2003) DNA methylation and mycoplasma genomes. *J Mol Evol* 57:S21–S28
- Xia X (2005) Mutation and selection on the anticodon of tRNA genes in vertebrate mitochondrial genomes. *Gene* 345:13–20
- Xia X (2007a) Bioinformatics and the cell: modern computational approaches in genomics, proteomics and transcriptomics. Springer, New York
- Xia X (2007b) An improved implementation of codon adaptation index. *Evol Bioinform* 3:53–58
- Xia X (2008) The cost of wobble translation in fungal mitochondrial genomes: integration of two traditional hypotheses. *BMC Evol Biol* 8:211
- Xia X (2009) Information-theoretic indices and an approximate significance test for testing the molecular clock hypothesis with genetic distances. *Mol Phylogenet Evol* 52:665–676
- Xia X (2012a) DNA replication and strand asymmetry in prokaryotic and mitochondrial genomes. *Current Genomics* 13:16–27
- Xia X (2012b) Rapid evolution of animal mitochondria. In: Singh RS, Xu J, Kulathinal RJ (eds) Evolution in the fast lane: rapidly evolving genes and genetic systems. Oxford University Press, Oxford
- Xia X (2013) Wobble hypothesis. In: Maloy S, Hughes K (eds) Brenner's encyclopedia of genetics. 2nd edition, Academic Press
- Xia X, Huang H, Carullo M, Betran E, Moriyama EN (2007) Conflict between translation initiation and elongation in vertebrate mitochondrial genomes. *PLoS ONE* 2:e227
- Xia X, Li WH (1998) What amino acid properties affect protein evolution? *J Mol Evol* 47:557–564
- Xia X, MacKay V, Yao X, Wu J, Miura F, Ito T, Morris DR (2011) Translation initiation: a regulatory role for Poly(A) tracts in front of the AUG codon in *Saccharomyces cerevisiae*. *Genetics* 189:469–478
- Xia X, Palidwor G (2005) Genomic adaptation to acidic environment: evidence from *Helicobacter pylori*. *Am Nat* 166:776–784
- Xia X, Wang HC, Xie Z, Carullo M, Huang H, Hickey DA (2006) Cytosine usage modulates the correlation between CDS length and CG content in prokaryotic genomes. *Mol Biol Evol* 23:1450–1454
- Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92:371–373

- Xia X, Yang Q (2011) A distance-based least-square method for dating speciation events. *Mol Phylogenet Evol* 59:342–353
- Xia X, Yang Q (2013) Cenancestor. In: Maloy S, Hughes K (eds) *Brenner's encyclopedia of genetics*. 2nd edition, Academic Press
- Xia X, Yuen KY (2005) Differential selection and mutation between dsDNA and ssDNA phages shape the evolution of their genomic AT percentage. *BMC Genet* 6:20
- Xia XH, Wei T, Xie Z, Danchin A (2002) Genomic changes in nucleotide and dinucleotide frequencies in *Pasteurella multocida* cultured under high temperature. *Genetics* 161:1385–1394
- Yang MY, Bowmaker M, Reyes A, Vergani L, Angeli P, Gringeri E, Jacobs HT, Holt IJ (2002) Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell* 111:495–505
- Yasukawa T, Yang M-Y, Jacobs HT, Holt IJ (2005) A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Mol Cell* 18:651
- Yu Q, Chen D, König R, Mariani R, Unutmaz D, Landau NR (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J Biol Chem* 279:53379–53386
- Zhang D, Xiong H, Shan J, Xia X, Trudeau V (2008) Functional insight into Maelstrom in the germline piRNA pathway: a unique domain homologous to the DnaQ-H 3'-5' exonuclease, its lineage-specific expansion/loss and evolutionarily active site switch. *Biology Direct* 3:48
- Zilberstein D, Agmon V, Schuldiner S, Padan E (1982) The sodium/proton antiporter is part of the pH homeostasis mechanism in *Escherichia coli*. *J Biol Chem* 257:3687–3691
- Zilberstein D, Padan E, Schuldiner S (1980) A single locus in *Escherichia coli* governs growth in alkaline pH and on carbon sources whose transport is sodium dependent. *FEBS Lett* 116:177–180

Index

A

Acid resistance, 8
Acid-adaptation hypothesis (AAH), 5, 7
Acid-resistant, 3, 4, 36, 37
Adaptation, 4–7, 9, 10, 13, 20, 24, 37
Amino acid
 arginine, 4
 positively charged, 4
Ancestry, 55
APOBEC3, 9
Arginine, 4, 8, 15, 16, 31, 32

B

Bootscan, 49, 51, 52

C

Caenorhabditis elegans, 16, 37, 42
Cephalochordates, 15
C-minimization hypothesis, 6
Coancestry, 13
Codon usage, 8–11, 13, 15, 20, 31
Codon-anticodon adaptation, 9, 10, 13, 20
Comparative method, 23, 25, 32, 34, 42, 43
 continuous variable
 discrete variable
Compatibility, 49, 54–56
Compatibility matrix, 49, 54–56
Convergence, 13
CpG, 11, 12, 20
 deficiency, 11, 12, 20
Cytoplasmic pH, 3, 6–9, 37

D

DAMBE, 8, 33, 42, 45, 50–52

Deamination, 5, 11, 19
Drosophila melanogaster, 37, 42

E

ERG, 37, 42
Escherichia coli, 2, 5–7, 18, 19, 34, 37
Exaptation, 4, 5

F

False discovery rate (FDR), 23, 43–45, 54
Familywise error rate (FWER), 43–45

G

GC skew, 17–19
Genome duplication, 46

H

Helicobacter hepaticus, 8, 9, 20
Helicobacter pylori, 3–9, 20, 35, 37
HIV-1, 9, 10, 20, 21, 50–52
Homology, 1, 9, 12, 13, 22, 32, 42, 46, 51
HTLV-1, 9, 10, 20

I

Internal pH, 6
Inversion distance, 19, 46
IRA2, 2
Isoelectric point (pI), 4–9, 20–22
Isoleucine, 9, 15

K

KAL153, 50–53

L

Laccase, 21, 22
 Lancelets, 15
 Lateral gene transfer (LGT), 1, 33–36
 Likelihood ratio test, 39, 45
 Lysine, 4–6, 8

M

Maelstrom, 42
 Membrane proteins, 4, 5, 7
 Methylation, 11, 12, 20
 Methyltransferase, 11, 12, 20, 37
 Minimal genome, 42
 Mutation, 1, 9, 10, 12, 19, 20, 30–32, 54
 AT-biased, 32
Mycoplasma genitalium, 11, 12
Mycoplasma pneumoniae, 11, 12
Mycoplasma pulmonis, 12, 15

N

NF1, 2

O

Orthology, 13, 21, 42, 46

P

Palindrome, 47
 Paralogy, 13, 36
Pasteurella multocida, 24
 Phylogenetic incongruence test, 23, 33–36,
 49, 51, 52, 54
 Phylogeny, 11–13, 24–26, 28, 32–34,
 38, 39, 42

poly(A), 22

Preadaptation, 4, 6

Precipitation avoidance hypothesis (PAH), 7–9

Pruning algorithm, 9, 40, 41

R

Reverse transcriptase, 9

S

Shigella flexneri, 2

Simplot, 49–52

Strand asymmetry, 1, 17, 19, 20

Strand-coupled bidirectional replication
 (mtDNA), 19

Strand-displacement model (mtDNA), 18, 19

Synteny, 46, 47

T

tRNA pool, 9, 10, 15

Tunicates, 13, 15

U

Urea, 3, 4, 37

Urease, 3, 4, 35, 37, 38

Urochordates, 15

W

Wobble hypothesis, 15–17
 expanded, 15–17