

SUDHIR KUMAR

COMPARATIVE GENOMICS AND BIOINFORMATICS

ADDRESS

Center for Evolutionary Functional Genomics
Biodesign Institute A-240
727 East Tyler Street
Arizona State University
Tempe, AZ 85287-5301
www.kumarlab.net

PHONE: (623) 225-5230 (mobile)
(480) 727-6949 (Office/Assistant)
FAX: (480) 727-6947
E-MAIL: s.kumar@asu.edu
www.biodesign.asu.edu

EDUCATION

	DEGREE	YEAR	DISCIPLINE
Birla Institute of Tech. & Sci. (India)	B. Engg.	1990	Electrical & Electronics Engg*
Birla Institute of Tech. & Sci. (India)	M. Sc.	1990	Biological Sciences*
Pennsylvania State University (USA)	Ph. D.	1996	Genetics (Masatoshi Nei)

*Concurrent degrees

PROFESSIONAL APPOINTMENTS

2006 – Professor, School of Life Sciences (SoLS), Arizona State University (ASU)
2003 – Director, Center for Evolutionary Functional Genomics, Biodesign Institute
2003 – 2004 Faculty Leader, Genomics, Evolution, and Bioinformatics, SoLS, ASU
2002 – 2006 Associate Professor, SoLS, ASU
1998 – 2002 Assistant Professor, Department of Biology, ASU
1996 – 1998 Postdoctoral Fellow (Nei), Department of Biology, Penn State University
1991 – 1996 Research Assistant (Nei), Department of Biology, Penn State University

AWARDS & HONORS

2009 CSE Honorary Professor, School of Computing and Informatics, ASU
2009 Fellow, American Association for the Advancement of Science (AAAS)
2007 – 2008 Japanese Society for Promotion of Science (JSPS) Visiting Fellowship Award, National Institutes of Genetics, Mishima, Japan (December-March)
2006 Exemplar Faculty (Promotion to Full Professor), ASU
2004 – Top-10 Most-Cited author, Web of Science (Computer Science)
2004 – 2006 Secretary (Elected), *Society for Molecular Biology and Evolution*
2000 Innovation Award in Functional Genomics, Burroughs-Wellcome Fund

RECOGNITIONS OF PUBLISHED WORKS

2009 H-index = 34; Total citations to date = 19,500+ (Google Scholar, July 2009)
2009 New Hot Paper in Computer Science (*Briefings in Bioinformatics*, 9:299-306), Essential Science Indicators (www.sciencewatch.com)
2007 – 2009 Current Classics in Computer Science (*Briefings in Bioinformatics* 5: 150-163), Essential Science Indicators
2006 Editor's Choice, SCIENCE, August 25, 2006 (*Molecular Biology & Evolution* 23:1946-1951)
2006 Hot Paper in Computer Science (*Briefings in Bioinformatics*, 5:150-163), Essential Science Indicators
2004 Hot Paper in Biology and in Computer Science (*Bioinformatics* 17:1244-1245), Web of Science, ISI
2000 Hot Paper in Biology (*Nature* 392: 917-920), Web of Science, ISI

MAJOR RESEARCH THEMES

Translational and fundamental investigations for (a) understanding evolutionary patterns of genomes and species, (b) predicting of disease-associated mutations in humans, and (c) large-scale synthetic analysis of gene expression patterns. A major focus of our work is on the bioinformatics development of software tools and knowledge-bases that accelerate biological discovery.

CURRENT RESEARCH GRANT SUPPORT (TOTAL COSTS FOR THE MOST RECENT AWARDS)AS THE PRINCIPAL INVESTIGATOR

- 2000 – 2010 *Comparative Molecular Sequence Analysis* (R01)
National Institute of Health (NHGRI)
- 2003 – 2010 *Computational Analysis of Gene Expression Pattern Images* (R01)
National Institute of Health (NHGRI)
- 2007 – 2010 *Re-engineering the MEGA Software Package* (R01)
National Institute of Health (NIGMS)

AS A CO-PRINCIPAL INVESTIGATOR OR CO-INVESTIGATOR

- 2006 – 2010 *Developing a Bioinformatic Database for Stoichioproteomics*
Co-principal Investigator (W. Fagan, PI)
National Science Foundation (DBI)
- 2006 – 2010 *Machine Learning Approaches for Biological Image Informatics*
Co-principal Investigator (J. Ye, PI)
National Science Foundation (IIS)
- 2008 – 2012 *Discovering the Hidden Proteome in the Human Genome*
Co-principal Investigator (J. Chaput, PI)
National Institutes of Health (EUREKA program)
- 2009 – 2010 *Team Approach to Translate Novel Biomarkers for Diabetes*
Co-Investigator (R. Nelson, PI)
National Institutes of Health (NIDDK)

COMPLETED RESEARCH GRANT SUPPORT

- 2000 – 2004 *Design of a Bioinformatic Database for Functional Evolutionary Footprints in Multigene Families*, Principal Investigator, National Science Foundation (DBI)
- 2000 – 2008 *Emerging Wildlife Diseases: Threats to Amphibian Biodiversity*, Co-Investigator, National Science Foundation; PI: J. Collins
- 2001 – 2004 *Development of an Evolutionary Timescale Database*, Co-Principal Investigator, National Science Foundation (DBI)
- 2001 – 2003 *A Computational Biosciences Professional Master's Program*, Participating Investigator, Sloan Foundation
- 2003 – 2006 *Computational Genomic Analysis to Identify and Dissect Functionally Important Mutations in Protein Sequences*, Burroughs-Wellcome Fund, Principal Investigator, USA
- 2004 – 2006 *LSAMP Bidesigned Bridges to the Doctorate*, Co-Principal Investigator, National Science Foundation (EHR)
- 2007 – 2008 *Bioinformatics of Assembling the Timescale of Life*, Principal Investigator, Science Foundation of Arizona

MEMBERSHIPS IN SCIENTIFIC AND PROFESSIONAL ORGANIZATIONS

American Association for the Advancement of Science
Arizona Arts, Sciences, and Technology Academy (founding fellow)
American Genetic Association
The Genetics Society of America
International Society of Computational Biologists
Society for Molecular Biology and Evolution
Society for the Study of Evolution

JOURNAL EDITORIAL AND WEBMASTER SERVICE

- 2005 – *Associate Editor*, Molecular Biology & Evolution
- 2005 – *Associate Editor*, Evolutionary Bioinformatics Online
- 2005 – *Editorial Board*, Genome Research
- 2004 – *Editorial Board*, Molecular and Developmental Evolution (J. Exp. Zool.)
- 2005 – 2006 *Board of Editors*, Gene: Functional Genomics
- 2004 – 2008 *Webmaster*, Society for Molecular Biology and Evolution (www.smbc.org)
- 1999 – 2007 *Webmaster*, American Genetic Association (www.theaga.org)

1999 – 2005 Associate Editor, Journal of Heredity

CONFERENCE/SYMPOSIA/OUTREACH ORGANIZED

- 2007 Hosted FlyBase outreach meeting at ASU (March)
- 2007 Member, Program Committee, International Conference on Computational Phylogenetics and Molecular Systematics, Moscow State University, Russia (November 16 – 19)
- 2006 Host, Annual meeting of the *Society for Molecular Biology and Evolution*, Arizona State University, Tempe, Arizona (May 24 – 28)
- 2005 Organizer, Genome Database Workshop @ *National Evolutionary Synthesis Center*, Wilmington Beach, North Carolina, (May 31 – Jun 3)
- 2004 Symposium on Evolutionary and Population Genomics in the *Future of Statistics Conference at Indian School of Business*, Hyderabad, India (Dec. 29 – Jan. 1)

PROFESSIONAL AFFILIATIONS

- 2007 – Guest Professor, Center for Computational and Evolutionary Biology (CCEB), Institute of Zoology at Beijing, China
- 2002 – Affiliate Professor, Computer Science & Engineering, ASU
- 2000 – Member, NASA Astrobiology/Evolutionary Genomics Focus Group
- 1998 – Associate Member, Penn State Astrobiology Research Center
- 2005 – 2007 Affiliate Professor, Department of Biomedical Informatics, ASU
- 2004 – 2007 Adjunct Senior Investigator, Translational Genomics Research Institute (TGen), Phoenix, Arizona
- 1998 – 2005 Member, Genetics Program, ASU

ORGANIZATIONAL REVIEWER & CONSULTANT

- 2006 Consultant, Advisory Committee on *Evolutionary Bioinformatics*, University of South Dakota, USA
- 2006 – 2007 Consultant, Amerigenics, Inc., USA
- 2005 Member, five-year review committee, Institute for Genomics and Bioinformatics (IGB) @ University of California, Irvine
- 2002 – 2003 Consultant, Pharmacia Corporation, USA

ADVISORY BOARDS / WORKING GROUPS

- 2007 FlyBase Advisory Group
- 2006 – Standing Member, BioData Management and Analysis (BDMA) Study Section, National Institutes of Health, USA
- 2006 – 2009 Member, Informatics Advisory Committee, National Center for Evolutionary Synthesis (NESCent), USA
- 2006 – 2008 Member, Working Group: Evolutionary Informatics: Supporting Interoperability in Evolutionary Analysis, NESCent, USA

PUBLICATIONS/PRODUCTS

DISCOVERY TOOLS AND KNOWLEDGE-BASES

1. Kumar S, Tamura K & Nei M (1994) *MEGA: Molecular Evolutionary Genetics Analysis* software for microcomputers. ***Computer Applications in Biosciences*** 10:189–191.
2. Kumar S (1995) *PhylTest: A Program for Testing Phylogenetic Hypotheses*. Pennsylvania State University, University Park.
3. Kumar S, Tamura K, Jakobsen IB & Nei M (2001) *MEGA2: Molecular Evolutionary Genetics Analysis* software. ***Bioinformatics*** 17:1244–1245.
4. Kumar S, Tamura K & Nei M (2004) *MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment*. ***Briefings in Bioinformatics*** 5:150–163.

5. Hedges SB, Dudley J & Kumar S (2006) TimeTree: a public knowledge-base of divergence times among organisms. **Bioinformatics** 22: 2971–2972.
6. Kumar S and the FlyExpress Consortium (2006/2009) FlyExpress: A knowledge-base of spatiotemporal expression patterns at a genomic-scale in the fruit-fly embryogenesis. Arizona State University, Tempe, Arizona. (www.flyexpress.net)
7. Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology & Evolution** 24:1596–1599.
8. Kumar S, Van Emden B, Acquisti C, Fagan WF, Elser JJ (2008) GRASP: Genomic Resource Access for Stoichioproteomics. Arizona State University, Tempe, Arizona. (www.graspdb.net)
9. Hedges SB and Kumar S (2009) TimeTree 2: An expanded public knowledge-base of divergence times among organisms. Pennsylvania and Arizona State Universities, USA. (www.timetree.org)

SYNTHESIS AND MONOGRAPHS

10. Nei M & Kumar S (2000) **Molecular Evolution and Phylogenetics**. Oxford University Press, New York (333 pp). (Translated in Chinese, Japanese, and Russian)
11. Gerber AS, Loggins R, Kumar S & Dowling TE (2001) Does non-neutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? **Annual Review of Genetics** 35:539–566.
12. Hedges SB & Kumar S (2002) Vertebrate genomes compared. **Science** 297:1283–1285.
13. Hedges SB & Kumar S (2003) Genomic clocks and evolutionary timescales. **Trends in Genetics** 19:200–206.
14. Hedges SB & Kumar S (2004) Precision of molecular time estimates. **Trends in Genetics** 20:242–247.
15. Kumar S (2005) Molecular clocks: four decades of evolution. **Nature Reviews Genetics** 6:654–662.
16. Kumar S & Hedges SB (2005) Pushing back the expansion of introns in animal genomes. **Cell** 123:1182–1184.
17. Filipski A & Kumar S (2005) Comparative Genomics in Eukaryotes. In **The Evolution of the Genome** edited by Ryan TG. Elsevier, San Diego, pp 521–583.
18. Kumar S & Filipski A (2007) Multiple sequence alignment: in pursuit of homologous DNA positions. **Genome Research** 17:127–135.
19. Kumar S & Dudley J (2007) Bioinformatics software for biologists in the genomics era. **Bioinformatics** 23:1713–1717.
20. Filipski A, Prohaska S & Kumar S (2007) Detecting molecular signatures of adaptive evolution. In **Evolutionary Genomics and Proteomics** edited by Pagel M & Pomiankowski A. Sinauer Associates, Massachusetts, pp 241–254.
21. Kumar S, Dudley J, Nei M & Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. **Briefings in Bioinformatics** 9:299–306.
22. Hedges SB & Kumar S (2009) **The Timetree of Life**. Oxford University Press, New York (550 pp; edited volume with 81 contributions).
23. Hedges SB & Kumar S (2009) Discovering the Timetree of Life. In **The Timetree of Life** edited by Hedges SB & Kumar S. Oxford University Press, New York, pp 3–18.

RESEARCH ARTICLES

24. Hedges SB, Kumar S, Tamura K & Stoneking M (1992) Human origins and analysis of mitochondrial DNA sequences. **Science** 255:737–739.
25. Kumar S, Tamura K & Nei M (1993) **A Guide to Molecular Evolutionary Genetics Analysis Program for Microcomputers**, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA (140 pp; >2500 printed manuals distributed).

26. Rzhetsky A, Kumar S & Nei M (1995) Four-cluster analysis: a simple method to test phylogenetic hypotheses. *Molecular Biology & Evolution* 12:163–167.
27. Yang Z, Kumar S & Nei M (1995) A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* 141:1641–1650.
28. Winnepenninckx W, Backeljau T, Mackey LY, Brooks JM, De-Wachter R, Kumar S & Garey JR (1995) 18S rRNA data indicate that Aschelminthes are polyphyletic in origin and consist of at least three distinct clades. *Molecular Biology & Evolution* 12:1132–1137.
29. Hedges SB, Parker PH, Sibley CG & Kumar S (1996) Continental breakup and the ordinal diversification of birds and mammals. *Nature* 381:226–229.
30. Kumar S (1996) A stepwise algorithm for finding minimum evolution trees. *Molecular Biology & Evolution* 13:584–593.
31. Kumar S (1996) Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. *Genetics* 143:537–548.
32. Kumar S, Balczarek KA & Lai Z-C (1996) Evolution of the *hedgehog* gene family. *Genetics* 142:965–972.
33. Kumar S & Rzhetsky A (1996) Evolutionary relationships of eukaryotic kingdoms. *Journal of Molecular Evolution* 42:183–193.
34. Yang Z & Kumar S (1996) Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. *Molecular Biology & Evolution* 13:650–659.
35. Balczarek KA, Lai Z-C & Kumar S (1997) Evolution and functional diversification of the Paired box (*Pax*) DNA-binding domains. *Molecular Biology & Evolution* 14:829–842.
36. Zhang J & Kumar S (1997) Detection of convergent and parallel evolution at the amino acid sequence level. *Molecular Biology & Evolution* 14:527–536.
37. Zhang J, Kumar S & Nei M (1997) Small-sample tests of episodic adaptive evolution: A case study of primate lysozymes. *Molecular Biology & Evolution* 14:1335–1338.
38. Yeager M, Kumar S & Hughes AL (1997) Sequence convergence in the peptide-binding region of primate and rodent MHC class *Ib* molecules. *Molecular Biology & Evolution* 14:1035–1041.
39. Leitner TL, Kumar S & Albert J (1997) Tempo and mode of nucleotide substitutions in *gag* and *env* gene fragments in Human Immunodeficiency Virus Type 1 populations with a known transmission history. *Journal of Virology* 71:4761–4770.
40. Kumar S & Hedges SB (1998) A molecular timescale for vertebrate evolution. *Nature* 392:917–920.
41. Nei M, Kumar S & Takahashi K (1998) The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. *Proceedings of the National Academy of Sciences (USA)* 95:12390–12397.
42. Hedges SB & Kumar S (1999) Divergence times of eutherian mammals. *Science* 285:2031a.
43. Newfeld SJ, Wisotzkey RG & Kumar S (1999) Molecular evolution of a development pathway: Phylogenetic analyses of transforming growth factor- β family ligands, receptors, and Smad signal transducers. *Genetics* 152:783–795.
44. Wang Y-C, Kumar S & Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals, and fungi. *Proceedings of the Royal Society, London. B* 266:163–171.
45. O'Brien S, Eisenberg JF, Miyamoto M, Hedges SB, Kumar S, & Wilson DE (1999) Genome Maps 10. Comparative Genomics. Mammalian radiations. Wall Chart. *Science* 286:463–478.
46. Kumar S, Mitnik C, Valente G & Floyd-Smith G (2000) Expansion and molecular evolution of the interferon-induced 2'-5' oligoadenylate synthetase gene family. *Molecular Biology & Evolution* 17:738–750.

47. Kumar S, Hedrick P, Dowling T, & Stoneking M (2000) Questioning evidence for recombination in human mitochondrial DNA. **Science** 288:1931a.
48. Kumar S & Gadagkar SR (2000) Efficiency of the neighbor-joining method in reconstructing deep and shallow evolutionary relationships in large phylogenies. **Journal of Molecular Evolution** 51:544–553.
49. Purdom PW, Bradford PG, Tamura K & Kumar S (2000) Single column discrepancy and dynamic max-mini optimizations for quickly finding the most parsimonious evolutionary trees. **Bioinformatics** 16:140–151.
50. Kumar S, Gadagkar SR, Filipinski A & Gu X (2001) Determination of the number of conserved chromosomal segments between species. **Genetics** 157:1387–1395.
51. Kumar S & Gadagkar SR (2001) Disparity Index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. **Genetics** 158:1321–1327.
52. Kumar S & Panchanathan S (2001) Elucidating gene interaction networks based on gene expression pattern image analysis. **Proceedings of the International Conference on Biomedical Engineering** 5A:232–234.
53. Rosenberg MS & Kumar S (2001) Traditional phylogenetic reconstruction methods reconstruct shallow and deep evolutionary relationships equally well. **Molecular Biology & Evolution** 18:1823–1827.
54. Rosenberg MS & Kumar S (2001) Incomplete taxon sampling is not a problem for phylogenetic inference. **Proceedings of the National Academy of Sciences (USA)** 98:10751–10756.
55. Miller MP & Kumar S (2001) Understanding human disease mutations through the use of interspecific genetic variation. **Human Molecular Genetics** 10: 2319–2328.
56. Hedrick P & Kumar S (2001) Mutation and linkage disequilibrium in human mtDNA. **European Journal of Human Genetics** 9:969–972.
57. Hedges SB, Chen H, Kumar S, Wang DY-C, Thompson AS, & Watanabe H (2001) A genomic timescale for the origin of eukaryotes. **BMC Evolutionary Biology** 1:4 (10 pp).
58. Jayaraman K, Panchanathan S, & Kumar S (2001) Classification and indexing of gene expression images. **Proceedings of Society of Photo-optical Instrumentation Engineers** 4472:471–481.
59. Kumar S & Subramanian S (2002) Mutation rates in mammalian genomes. **Proceedings of the National Academy of Sciences (USA)** 99:803–808.
60. Kumar S, Jayaraman K, Panchanathan S, Gurunathan R, Marti-Subirana A & Newfield SJ (2002) BEST: A novel computational approach for comparing gene expression patterns from early stages of *Drosophila melanogaster* development. **Genetics** 162:2037–2047.
61. Tamura K & Kumar S (2002) Evolutionary distance estimation under heterogeneous substitution pattern among lineages. **Molecular Biology & Evolution** 19:1727–1736.
62. Jiang Z, Melville JS, Cao H, Kumar S, Filipinski A & Verrinder Gibbins AM (2002) Measuring conservation of contiguous sets of autosomal markers on bovine and porcine genomes in relation to the map of the human genome. **Genome** 45:769–776.
63. Subramanian S & Kumar S (2003) Neutral substitutions occur as a faster rate in exons than in noncoding DNA in primate genomes. **Genome Research** 13:838–844.
64. Rosenberg MS & Kumar S (2003) Taxon sampling, bioinformatics, and phylogenomics. **Systematic Biology** 52:119–124.
65. Rosenberg MS, Subramanian S & Kumar S (2003) Patterns of transitional mutation biases within and among mammalian genomes. **Molecular Biology & Evolution** 20:988–993.
66. Rosenberg MS & Kumar S (2003) Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference. **Molecular Biology & Evolution** 20:610–621.
67. Miller MP, Parker JD, Rissing SW, & Kumar S (2003) Quantifying the intragenic distribution of human disease mutations. **Annals of Human Genetics** 67:567–579.

68. Jancovich J, Mao J, Chinchar VG, Wyatt C, Case S, Kumar S, Valente G, Subramanian S, Davidson EW, Collins JP & Jacobs BL (2003) Genomic sequence of a ranavirus (family Iridoviridae) associated with salamander mortalities in North America. ***Virology*** 316:90–103.
69. Tamura K, Subramanian S & Kumar S (2004) Temporal patterns of fruit fly evolution revealed by mutation clocks. ***Molecular Biology and Evolution*** 21:36–44.
70. Briscoe A, Gaur C & Kumar S (2004) The spectrum of human rhodopsin disease mutations through the lens of interspecific variation. ***Gene*** 332:107–118.
71. Tamura K, Masatoshi Nei, & Kumar S (2004) Prospects for inferring very large phylogenies using the neighbor-joining method. ***Proceedings of the National Academy of Sciences (USA)*** 101:11030–11035.
72. Subramanian S & Kumar S (2004) Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome. ***Genetics*** 168:373–381.
73. Gurunathan R, Van Emden B, Panchanathan S & Kumar S (2004) Identifying spatially similar gene expression patterns in early stage fruit fly embryo images: binary feature versus invariant moment digital representations. ***BMC Bioinformatics*** 5:202 (13 pp).
74. Gargesha M, Antin P, Van Emden B, Panchanathan S & Kumar S (2004) Image registration and similarity computation for chicken gene expression patterns. ***Genomic Signal Processing and Statistics (GENSIPS)***, IEEE Signal Processing Society (4 pp).
75. Kumar S, Filipski A, Swarna V, Walker A & Hedges SB (2005) Placing confidence limits on the molecular age of the human-chimpanzee divergence. ***Proceedings of the National Academy of Sciences (USA)*** 102:18842–18847.
76. Gadagkar SR, Rosenberg MS & Kumar S (2005) Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. ***Journal of Experimental Zoology (Molecular & Developmental Evolution)*** 304B:64–74.
77. Gadagkar SR & Kumar S (2005) Maximum likelihood outperforms maximum parsimony even when evolutionary rates are heterotachous. ***Molecular Biology and Evolution*** 22:2139–2141.
78. Gargesha M, Yang J, Van Emden B, Panchanathan S & Kumar S (2005) Automatic annotation techniques for gene expression images of the fruit fly embryo. ***Proceedings of Society of Photo-optical Instrumentation Engineers*** 5960:576–583.
79. Hedges SB, Kumar S & van Tuinen M (2006) Constraining fossil calibrations for molecular clocks. ***BioEssays*** 28:770–771.
80. Elser JJ, Fagan WF, Subramanian S & Kumar S (2006) Signatures of ecological resource availability in the animal and plant proteomes. ***Molecular Biology & Evolution*** 23:1946–1951.
81. Subramanian S & Kumar S (2006) Higher intensity of natural selection on >90% of the human genes revealed by the intrinsic replacement mutation rates. ***Molecular Biology & Evolution*** 23:2283–2287.
82. Subramanian S & Kumar S (2006) Evolutionary anatomies of positions and types of disease-associated and neutral amino acid mutations in the human genome. ***BMC Genomics*** 7:306 (9 pp).
83. Ye J, Chen J, Li Q & Kumar S (2006) Classification of *Drosophila* embryonic developmental stage range based on gene expression pattern images. ***Computational Systems Bioinformatics Conference*** 4: 293–298.
84. Xia X & Kumar S (2006) Codon-based detection of positive selection can be biased by heterogeneous distribution of polar amino acids along protein sequences. ***Computational Systems Bioinformatics Conference*** 4: 335–340.
85. Colbourn CJ & Kumar S (2007) Lower bounds on multiple sequence alignment using exact 3-way alignment. ***BMC Bioinformatics*** 8:140 (8 pp).
86. Zhang Y, Sturgill D, Parisi M, Kumar S & Oliver B (2007) Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. ***Nature*** 450:233–237.

87. Clark AG, ..., Kumar S, ..., Drosophila 12 Genomes Consortium (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. **Nature** 450:203–218.
88. Acquisti C, Poste G, Curtiss D & Kumar S (2007) Nullomers: really a matter of natural selection? **PLoS ONE** 2 (3 pp).
89. Ye J, Chen J, Janardan R & Kumar S (2008) Developmental stage annotation of *Drosophila* gene expression pattern images via an entire solution path for LDA. **ACM Transactions on Knowledge Discovery from Data** 2:1–21.
90. Ji S, Sun L, Jin R, Kumar S, & Ye J (2008) Automated annotation of *Drosophila* gene expression patterns using a controlled vocabulary. **Bioinformatics** 24:1881–1888.
91. Kumar S, Suleski M, Markov GJ, Lawrence S, Marco A & Filipinski AJ (2009) Positional conservation and amino acids shape the correct diagnosis and population frequencies of benign and damaging personal amino acid mutations. **Genome Research** (doi:10.1101/gr.091991.109).
92. Suzuki Y, Gojobori T & Kumar S (2009) Methods for incorporating the hypermutability of CpG dinucleotides in detecting natural selection operating at the amino acid sequence level. **Molecular Biology & Evolution** (doi:10.1093/molbev/msp133).
93. Marco A, Konikoff C, Karr TL & Kumar S (2009) Relationship between gene co-expression and sharing of transcription factor binding sites in *Drosophila melanogaster*. **Bioinformatics** (doi:10.1093/bioinformatics/btp462).
94. Acquisti C, Elser JJ & Kumar S (2009) Nitrogen-limitation shapes the DNA composition of plant genomes. **Molecular Biology & Evolution** 26:953–956.
95. Acquisti C, Kumar S & Elser JJ (2009) Signatures of nitrogen limitation in the elemental composition of the proteins involved in the metabolic apparatus. **Proceedings of the Royal Society, London, B** (doi:10.1098/rspb.2008.1960).
96. Goering LM, Hunt PK, Heighington C, Busick C, Pennings P, Hermisson J, Kumar S & Gibson G (2009) Association of orthodenticle with natural variation for early embryonic patterning in *Drosophila melanogaster*. **Journal of Experimental Zoology (Molecular & Developmental Evolution)** doi:10.1002/jez.b.21299.
97. Kirill VM, Konstantinova AV, Nikitin MA, Troshin PV, Rusin LY, Lyubetsky VA, Panchin YV, Mylnikov AP, Moroz LL, Kumar S & Aleoshin VV (2009) The origin of Metazoa: a transition from temporal to spatial cell differentiation. **BioEssays** 31:758-768.
98. Horiike T, Miyata D, Hamada K, Saruhashi S, Shinozawa T, Kumar S, Chakraborty R, Komiyama T & Tateno Y (2009) Phylogenetic construction of 17 bacterial phyla by new method and carefully selected orthologs. **Gene** 429:59–64.
99. Ji S, Li Y-X, Zhou Z-H, Kumar S & Ye J (2009) A bag-of-words approach for *Drosophila* gene expression pattern annotation. **BMC Bioinformatics** 10:119 (16 pp).
100. Li Y-X, Ji S, Kumar S, Ye J & Zhou Z-H (2009) *Drosophila* gene expression pattern annotation by multi-instance multi-label learning. International Joint Conference on Artificial Intelligence (IJCAI) (in press).
101. Ji S, Yuan L, Li Y-X, Zhou Z-H, Kumar S & Ye J (2009) *Drosophila* gene expression pattern annotation using sparse features and term-term interactions. **ACM SIGKDD International Conference on Knowledge Discovery and Data Mining** (in press).

Encyclopedia, Dictionaries, and Book Reviews

102. Kumar S (2000) A review of the book *Genomes* by TA Brown. **The Quarterly Review of Biology** 75:316–317.
103. Kumar S & Filipinski A (2001 and 2008) Molecular Phylogeny Reconstruction. **Encyclopedia of Life Sciences** Macmillan Reference Ltd, Oxford, UK. (www.els.net).
104. Kumar S & Filipinski A (2001) Molecular Clock Testing. **Encyclopedia of Life Sciences**, Macmillan Reference Ltd, Oxford, UK. (www.els.net).
105. Kumar S & Newfeld SJ (2002) A review of the book *Modern Genetic Analysis: Integrating Genes and Genomes* (second edition) by Griffiths AJF, Gelbart WM, Lewontin RC & Miller JH. **The Quarterly Review of Biology** 77:456–457.

106. Rawls A & Kumar S (2002) A review of the book *Genomic Regulatory Systems: Development and Evolution* by E. H. Davidson. **The Quarterly Review of Biology** 77:456.
107. Lorson C & Kumar S (2003) A review of the book *Genomes (second edition)* by T.A. Brown. **The Quarterly Review of Biology** 78:225.
108. Kumar S (2003) *MacTrees made easy*, a review of the book *Phylogenetic trees made easy: a how-to-manual for molecular biologists* by Hall BG. **Molecular Evolution & Phylogenetics** 27:165-167.
109. Kumar S & Filipinski A (2004) Phylogenetic Analysis. **Dictionary of Bioinformatics and Computational Biology** edited by Hancock J & Zvelebil M, Wiley-Liss, New York.
110. Kumar S & Filipinski A (2005) Reconstructing Vertebrate Phylogeny. **Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics** edited by Subramaniam S. John Wiley & Sons, New York.
111. Kumar S with many authors (2007) Sequence Assembly and Alignment Tech Guide. **Genome Technology** (10 pp).
112. Ye J, Janardan R & Kumar S (2008) Biological Image Analysis via Matrix Approximation, **Encyclopedia of Data Warehousing and Mining** (2nd Ed.) edited by John Wang, Idea Group, Inc., Pennsylvania (Pp. 166-170) .

INVITED PRESENTATIONS

(Titles included for symposia, workshops, and special presentations)

- 2009 Teaching Workshop @ Centers for Disease Control (CDC), Atlanta (July 13/14)
Title: Molecular Evolution and Phylogenetics
- 2009 Chalk Talk, Physics Department, ASU
- 2009 Spirit of Senses Group, Phoenix, Arizona
- 2009 University of Cologne, Germany
- 2009 Quantitative Expression Analysis workshop @ Drosophila Research Conference, Chicago, Illinois
Title: Challenges in Integrative Analysis of High-Throughput Datasets
- 2008 University of North Carolina, Charlotte, NC
- 2008 Washington University, St. Louis, MO
- 2008 Symposium on *Evolutionary Bioinformatics*, Keynote address, Lava Springs, Idaho State University (Oct 3-5)
Title: Software tools and Knowledge-bases for discoveries from genomes to phenotypes
- 2008 Japan Biological Information Research Center (JBIRC), Tokyo, Japan.
- 2008 Symposium on *New Insight of Genome Evolution into Fundamental Activities of Life*, National Institute of Genetics (NIG) and the Tokyo Institute of Technology (TIT), Japan
- 2008 Global Center for Excellence, Hokkaido University, Sapporo, Japan.
- 2008 Discussion leader, *Computational and Statistical Advances*, Gordon Conference in Molecular Evolution @ Ventura, California (Feb 3-8, 2008)
- 2007 SOLUR Program, Arizona State University
- 2007 Department of Biomedical Informatics, Arizona State University
- 2007 FlyExpress: The Next Steps. FlyBase Advisory Group @ Harvard University, Boston, MA (Sept 15-18)
- 2007 Keynote Speaker, Ohio Collaborative Conference on Bioinformatics (OCCBIO) @ Miami University, Oxford, Ohio
Title: Evolutionary anatomies of disease mutations in proteins
- 2006 EMBO workshop on "Human Evolution and Disease" @ Center for Cellular and Molecular Biology, Hyderabad, India
Title: Evolutionary anatomies of the locations and types of disease-associated nonsynonymous mutations in the human genome
- 2006 Department of Biological Sciences, University of Idaho, Moscow, ID
- 2006 Microbiology Department, Montana State University, Bozeman, MT
- 2005 ASU Emeritus Faculty Association Luncheon, Tempe, AZ

- Title: Evolutionary Functional Genomics in the Biodesign Institute: Taking a Computational Approach to Understanding and Harnessing the Biological Design*
- 2005 Symposium on *Molecular Evolution @ Moscow Conference on Computational Molecular Biology*, Moscow State University, Russia
Title: Placing Confidence Limits on the Molecular Age of the Human-Chimpanzee Divergence
- 2005 Symposium on Systems Biology @ Moscow Conference on Computational Molecular Biology, Moscow State University, Russia
Title: Automating Discovery of Gene Interactions by in silico Analysis of in situ Gene Expression Patterns in Fruit Fly Embryos
- 2005 Special presentation to the *Panel on Chemical Imaging*, National Academies (USA), Washington DC
Title: In silico Analysis and Management of Fruit Fly Gene Expression Pattern Images
- 2004 Symposium on *Evolutionary and Population Genomics @ Future of Statistics Conference*, Hyderabad, India
Title: Genomic Bounds on the Timing of the Evolutionary Divergence of Humans and Chimpanzees
- 2004 Symposium on *Advances in Methods for Estimating Species Divergence Dates Using Molecular Data @ International Congress of Zoology*, Beijing, China
Title: Genomic Timescales: Precision & Robustness
- 2004 Workshop on *Molecular Evolution and Systematics*, Chinese Academy of Sciences, Beijing, China
Title: Distance and Parsimony Methods
- 2004 Symposium on *Molecular Phylogeny and Molecular Clocks @ Annual Meeting of the Society for Molecular Biology & Evolution*, Penn State University, University Park, PA
Title: Genomic Timescales: Precision & Robustness
- 2004 Hexapodium, Center for Insect Research, University of Arizona
- 2003 Techniques Workshop @ 44th Annual *Drosophila Research Conference*, Chicago, IL
Title: Basic Expression Search Tool (BEST): Computational Framework for in silico Analysis of in situ Hybridization Data
- 2003 *Comparative and Functional Genomics Workshop*, Sponsored by the *Wellcome Trust and Department of Energy*, Hinxton, Cambridgeshire, UK
Title: Patterns of Point Mutation and Protein Substitution Rates Revealed by Comparative Mammalian Genomics
- 2003 Annual Meeting of the *Society for Molecular Biology & Evolution*, Newport Beach, CA
- 2003 Department of Biology, Duke University, Durham, NC
- 2003 Bioinformatics Research Center, North Carolina State University, Raleigh, NC
- 2002 Symposium on *Evolutionary Genetics @ Annual meeting of the American Genetic Association*, Arizona State University, Tempe, Arizona
Title: Rates of Point Mutation in Mammals
- 2002 The 12th International Workshop on *Beyond the Identification of Transcribed Sequences: Functional, Evolutionary, and Expression Analysis* sponsored by *Department of Energy*, Washington, DC
Title: Building Fruit Fly Developmental Networks: in silico Approaches for analyzing in situ Gene Expression Patterns
- 2002 The 18th *International Symposium* in Conjunction with Award of the *International Prize for Biology*, Tokyo, Japan
Title: Building Fruit Fly Developmental Networks: in silico approaches for Analyzing in situ Gene Expression Patterns
- 2002 Department of Biology, Indiana University, Bloomington, Indiana
- 2002 Department of Computer Science, Arizona State University, Tempe, AZ
- 2002 Department of Biology, University of Michigan, Ann Arbor, MI
- 2001 International Workshop on *Population Genetics @ University of Montreal*, Montreal, Canada
Title: Estimating Neutral Evolutionary Rates in Mammals
- 2001 ASU President's Community Enrichment Program, Phoenix, Arizona

- 2001 Department of Biology, Ohio State University, Columbus, Ohio
 2001 Department of Biology, Hong Kong University, Hong Kong, China (2 lectures)
 2001 Department of Biology, Syracuse University, Syracuse, New York
 2001 Program in Ecology and Evolutionary Biology, University of Illinois, Urbana Champaign, Illinois
 2001 Department of Biology, Grand Canyon University, Phoenix, Arizona
 2001 Ecology & Evolutionary Biology Program/IGERT, Indiana University, Bloomington, Indiana
 2001 Department of Biology, Tokyo Metropolitan University, Tokyo, Japan
 2001 Biomedical Engineering, Indian Institute of Science, Bangalore, India
 2000 Department of Biology, Tokyo Metropolitan University, Tokyo, Japan
 2000 Graduate University for Advanced Studies, Hayama, Japan
 2000 Genetics Program, University of Arizona, Tucson, Arizona
 2000 Birla Institute of Technology & Sciences, Pilani, India
 2000 University of South Carolina, Columbia, South Carolina
 1999 Symposium on *Genomic Diversity @ Annual meeting of the American Genetic Association*, Pennsylvania State University, University Park, PA
 Title: Genomic Divergence between Species in terms of the Number of Chromosomal Rearrangements
 1999 Department of Biology, Arizona State University-West, Phoenix, Arizona
 1998 Department of Zoology and Genetics, Iowa State University, Ames, Iowa
 1997 Symposium on *Large Phylogenies @ Annual meeting of the Society for the Study of Systematic Biology*, University of Colorado, Boulder, Colorado
 Title: Inferring Large Phylogenies
 1997 National Cancer Institute, Frederick, Maryland
 1997 Department of Biology, Arizona State University, Tempe, Arizona

TEACHING, MENTORING, UNIVERSITY SERVICE

COURSES TAUGHT IN A CLASSROOM (Average rating in parentheses; SCALE: 1 = BEST; 4 = WORST)

2009	BIO 455 (Fall)	Introduction to Comparative Genomics	(30+ students)
2008	BIO 455 (1.2)	Introduction to Comparative Genomics	(10 students)
2007	BIO 494 (1.2)	Introduction to Comparative Genomics	(12 students)
2006	BIO 345 (1.4)	Organic Evolution	(180 students)
2005	BIO 345 (1.6)	Organic Evolution	(150 students)
2004	BIO 494 (1.3)	Introduction to Comparative Genomics	(11 students)
2003	BIO 345 (1.4)	Organic Evolution	(185 students)
2003	BIO 494 (1.2)	Introduction to Comparative Genomics	(12 students)
2001	BIO 445 (1.4)	Organic Evolution	(180 students)
2001	BIO 494 (1.3)	Computational Genomics	(7 students)
2000	BIO 445 (1.7)	Organic Evolution	(106 students)
2000	BIO 594 (1.3)	Molecular Evolutionary Genetics	(7 students)
1999	BIO 594 (1.5)	Molecular Evolutionary Genetics	(10 students)
1999	BIO 494 (1.2)	Advanced Evolution	(14 students)

POSTDOCTORAL FELLOWS MENTORED

2008 –	Yingqin Luo, Ph.D., Bioinformatics <i>Project: Microbiological Evolution</i>
2008 –	Fabia Battistuzzi, Ph.D., Biology/Astrobiology <i>Project: Computational Biology of the Tree of Life</i>
2006 –	Claudia Acquisti, Ph.D., Molecular Evolution <i>Project: Evolutionary Proteomics and Ecology</i>
2001 –	Alan Filipinski, Ph.D., Formal Language Theory, Michigan State <i>Project: Molecular Timescales and Phylogeny</i>

PAST POSTDOCTORAL FELLOWS

2008 – 2009	Antonio Marco-Castillo, Ph.D., Genetics, Spain
2005 – 2007	Bao Hong Shen, Ph.D., Computer Science, Ariz. State Univ.

2005 – 2007 Anup Som, Ph.D., Bioinformatics, Javadpur Univ., India
 2003 – 2005 Christine Kuslich, Ph.D., Biomedical Sciences—Genetics, Univ. of Hawaii
 2000 – 2006 Sankar Subramanian, Ph.D., Microbiology, IARI, India
 2003 – 2004 Araxi Urrutia, Ph.D., Evolutionary Genomics, Univ. of Bath, UK
 2000 – 2003 Michael S. Rosenberg, Ph.D., Ecology & Evolution, SUNY-Stony Brook
 2000 – 2002 Mark P. Miller, Ph.D., Evol & Conservation Genetics—Comp. Biol., NAU
 1998 – 2003 Sudhindra R. Gadagkar, Ph.D., Biology (Aquaculture Genetics), Canada

CHAIR/CO-CHAIR OF GRADUATE STUDENT COMMITTEES

2007 – Charlotte Konikoff, Molecular and Cellular Biology (Co-Chair)
 2007 – 2008 Siddarth Selvaraj, Professional MS in Computational Biosciences (Chair)
 2005 – 2006 Bindu Koshy, MS, Electrical Engineering (Chair)
 2005 – 2007 Vinod Swarna, MS, Biology (Chair)
 2005 – 2006 Hector Ramos, Professional MS in Computational Biosciences (Chair)
 2002 – 2006 Rajalakshmi Gurunathan, Ph.D., Computer Science (Chair)
 2002 – 2006 Madhusudana Gargasha, Ph.D., Electrical Engg (Co-Chair)
 2003 – 2005 Stephanie Rogers, Professional MS in Computational Biosciences (Chair)
 2003 – 2005 Vinod Swarna, Professional MS in Computational Biosciences (Chair)
 2003 – 2005 Hojoon Lee, Professional MS in Computational Biosciences (Chair)
 2003 – 2005 Xiaofen Liu, Professional MS in Computational Biosciences (Chair)
 2003 – 2004 Shubhra Gupta, Professional MS in Computational Biosciences (Chair)
 2002 – 2005 Jian Yang, MS, Bioinformatics/Computational Biology (Chair)
 2002 – 2004 Sandhya Durvasala, MS, Computer Science & Engg (Co-Chair)
 2001 – 2003 Patrick Kolb, MNS, Biology (Chair)
 2000 – 2002 Rekha Iyer, MS, Molecular & Cell Biology (Chair)
 1999 – 2001 Karthik Jayaraman, MS, Electrical Engineering (Co-Chair)

MEMBER OF GRADUATE STUDENT COMMITTEES

2007 – Takahiro Maruki, Ph.D., Biology
 2007 – Shuiwang Ji, Ph.D., Computer Science
 2007 – Jianhui Chen, Ph.D., Computer Science
 2003 – Hugo F. Gante, Ph.D., Biology
 2001 – 2007 Michael Schwemm, M.S., Biology
 1998 – 2005 Peter Unmack, Ph.D., Biology
 1998 – 2005 Evan Carson, Ph.D., Biology
 1999 – 2005 Carla Hurt, Ph.D., Biology
 1999 – 2003 Daniel Garrigan, Ph.D., Biology

UNDERGRADUATE/HIGH SCHOOL RESEARCH ASSISTANTS

2008 – Simon Lawrence, Nicolas Feddern, Adithya Rajan
 2007 – Michael Suleski, Nicholas Peterson, Daniel Peterson
 2006 – 2008 Christopher Busick, Bryan Sexton
 2007 Veronica Shi (High School), Biodesign Summer Intern; Asaria Jimenez, Melizabeth Santana, Yea Jin Ko, Ariana Rodriguez, Stephanie Negron, Thania Martinez, Wilda Rivera, Victor Correa, Liris Gonzalez, Jose Maldonado, German Velez, Cristina Rivera (All from Puerto Rico)
 2001 – 2003 Jacob Reidhead
 2000 – 2001 Veena Ganeshan
 2000 – 2000 Emily Davenport
 2000 – 2000 Diana Tlougan
 1998 – 2000 Heather Wiemann, Evolutionary of multigene family
 1998 – 1998 Candice White, DNA sequence analysis
 1999 – 1999 Roman Johnson, Evolutionary of multigene family

OTHER RESEARCH/DEVELOPMENT LAB PERSONNEL

2008 – Michael McCutchan, Database Administrator
 2008 – Jana McAlpin, Scientific Programmer

2007 – Revak Raj, Graduate Associate
 2006 – Siddarth Selvaraj, Graduate Associate
 2005 – Lin-Wei Wu, Scientific Programmer
 2003 – Wayne Parkhurst, Multimedia Specialist
 2004 – 2008 Bernard Van Emden, Faculty Research Associate
 2003 – 2005 Quan Nguyen, Computer Data Base Specialist
 2001 – 2007 Joel Dudley, Faculty Research Associate
 2005 – 2005 Vesna Djinovic, Programming Assistant
 2005 – 2006 Sean Dudley, Systems Administrator
 2005 – 2008 Ashly Ruttman, Graduate associate
 2004 – 2006 Dana Desonie, Scientific Writer
 2004 – 2005 Joseph Svitak, Programmer
 1999 – 2007 Graziela Valente, Lab Manager & Research Technician
 2002 – 2004 Renee Grothe, Graphic Artist
 2001 – 2002 Ben Timmerick, Graphics & Multimedia
 2000 – 2004 David Schwartz, Programmer

MAJOR UNIVERSITY COMMITTEES (*all at ASU*)

2008 – *Chair*, Biodesign Personnel Committee
 2008 – *Member*, Strategic Planning Committee, SoLS
 2007 – *Member*, Graduate Initiatives in Informatics Committee
 2007 – *Member*, Web Committee, SoLS
 2005 – *Member*, Promotion Committee in the Biodesign Institute
 2005 – *Member*, Steering Committee, Post-Baccalaureate Research Education Program in Biomedical Sciences (PREP) program (NIH-funded)
 2005 – 2007 *Member*, Personnel Committee, Dept. of Biomedical Informatics
 2004 – *Member*, Executive Committee, Biodesign Institute
 2004 – 2005 *Member*, Insect Neurogenomics Search (open rank)
 2004 – 2006 *Member*, Taskforce on Biomedical Informatics
 Taskforce is charged with the planning, design, and development of a new Department of Biomedical Informatics
 2003 – 2004 *Member*, Executive Committee, School of Life Sciences
 2003 – 2004 *Member*, SoLS Founding Director
 2001 *Representative* from ASU in 3 formation meetings of the International Genomics Consortium: Johns Hopkins Univ, Baltimore, MD; M. D. Anderson Cancer Center, Houston, TX; Sun HealthCare, Scottsdale, AZ
 2000 – *Member*, Executive Committee, Computational Biosciences Program
 2000 – 2001 *Member*, Information Technology Hot Team, Prop 301
 2000 – 2001 *Member*, Joint Bioengineering & Computer Science (open rank)
 2000 *Member*, Planning Committee, Academic degree program in Comp. Biosci.
 1999 *Member*, Molecular Genetics (Asst. Prof.)