

PAVEL A. PEVZNER

Ronald R. Taylor Chair Professor of Computer Science
Departments of Computer Science and Engineering, APM 3132,
University of California at San Diego, La Jolla, CA 92093-0114
E-mail: ppevzner@cs.ucsd.edu, Phone: (858) 822-4365
FAX: (858) 534-7029 Office : APM 4802

Education

B.S.	Moscow Technological Transport Institute	1979	Applied Mathematics
Ph.D.	Moscow Institute of Physics and Technology	1988	Mathematics&Physics

Professional Experience

2006-present	Director, Center for Algorithmic and Systems Biology at UCSD
2006-present	Howard Hughes Medical Institute Professor
2002-present	Adjunct Professor of Mathematics University of California San Diego, CA
2000-present	Ronald R. Taylor Chair Professor of Computer Science University of California San Diego, CA
1995- 2000	Professor, Departments of Mathematics, Computer Science, and Molecular Biology University of Southern California, CA
1993-1995	Member, Institute for Molecular Evolutionary Genetics and Biotechnology Institute, The Pennsylvania State University, PA
1992-1995	Associate Professor, Department of Computer Science, The Pennsylvania State University, PA
1990-1992	Postdoctoral Research Associate (with M.Waterman) and Lecturer Department of Mathematics, University of Southern California, CA
1985-1990	Junior Scientist, Scientist, Senior Scientist, Laboratory of Mathematical Methods, National Center for Biotechnology NIIGENETIKA, Moscow, Russia (formerly Institute of Genetics of Microorganisms VNIIGENETIKA)
1986-1988	Ph.D. Trainee, Institute of Physics and Technology, Moscow, Russia

Professional Activities

Editorial Boards

- Executive Editor, *Journal of Computational Biology*
- Editorial Board, *Bioinformatics*
- Editorial Board, *Discrete Applied Mathematics*
- Editorial Board, *Journal of Computer and System Sciences*

- Editorial Board, *Journal of Association for Computing Machinery (ACM)*
- Editorial Board, *Journal of Bioinformatics and Computational Biology*
- Steering Committee, IEEE/ACM Transactions on Computational Biology and Bioinformatics.
- BMC Bioinformatics
- Biology Direct
- Editorial Board, *Gene-COMBIS*, 1994-1997
- Co-editor (with S. Istrail and M. Waterman). MIT Press *Computational Molecular Biology* book series.
- Co-editor (with S. Istrail and M. Waterman). Springer -Verlag *Lecture Notes in Bioinformatics* series.

Guest Editor

- Guest Editor of the special volume “Computer Genetics” of *Bio Systems*, **30**, issues 1-3, 1993
- Guest Editor of the special volume “Combinatorial Methods in DNA Mapping and Sequencing”, *Journal of Computational Biology*, **2**, issue 2, 1995
- Guest Editor (with S.Istrail and R.Shamir) of the special series “Computational Molecular Biology”, *Discrete Applied Mathematics*. First volume: **71**, 1996, second volume: **88**,1998, third volume **104**, 2000, fourth volume, 2003.
- Guest Editor of the special RECOMB issues, *J. Comp. Biology*. First issue: **4**, 1997, second issue: **5**, 1998, third issue: **6**, 1999, fourth issue: **7**, 2000
- Guest Editor (with Dick Karp, Ming Li, and Ron Shamir) of the special Computational Molecular Biology issues *Journal of Computer and System Sciences*, 2003, 2005, 2007

Professional Societies.

- Co-founder and Chair of Steering Committee, Annual International Computational Molecular Biology Conference (RECOMB), 1996-present
- Board of Directors, International Society for Computational Biology, 1998-2001.

Conference Organization

- Co-organizer, Mini-symposium *Combinatorial methods for genome rearrangements*, Los-Angeles, California, March,18, 1994
- Member of directorate and session organizer. *First World Congress on Computational Medicine and Biotechnology*, Austin, Texas, April,24-29, 1994
- Member of international organizing committee and session chair, *Third International conference on Bioinformatics and Supercomputing*, Tallahassee, Florida, June 1-4, 1994
- Co-chair, Program committee, DIMACS Computational Molecular Biology Year. Workshop *DNA mapping and sequencing*, New Brunswick, New Jersey, October 3-6, 1994
- Member, Program committee, DIMACS Computational Molecular Biology Year. Workshop *DNA sequence alignment*, Princeton, New Jersey, November 8-10, 1994
- Member, Program Committee, 6th Annual Symposium *Combinatorial Pattern Matching*, Helsinki, Finland, May, 1995

- Member, Advisory Committee, *4th DIMACS International Algorithm Implementation Challenge*, New Brunswick, New Jersey, September, 11-13, 1994
- Member, Program Committee, DIMACS Computational Molecular Biology Year. Workshop *Gene recognition*, Philadelphia, Pennsylvania, October, 1995
- Session Chair, Mathematics and Molecular Biology IV, Santa Fe, New Mexico, November, 1995
- Member, Program Committee, 2nd Sandia Workshop on Computational Molecular Biology, Albuquerque, New Mexico, March, 1996
- Member, Program Committee, Conference of Computational Molecular Biology to honor the 50th anniversary of ENIAC, Princeton, New Jersey, May, 1996
- Member, Program Committee, 4th Israeli Symposium on Theory of Computing and Systems, Jerusalem, Israel, June, 1996
- Member, Program Committee, 7th Annual Symposium *Combinatorial Pattern Matching*, Laguna Beach, California, June, 1996
- Member, Program Committee, 3rd South American Workshop on String Processing, Recife, Brazil, August, 1996
- Member, Program Committee, 1st Annual International Computational Molecular Biology Conference (RECOMB 97), Santa Fe, NM, January, 1997
- American Mathematical Society meeting, Computational Biology Section, Corvallis, Oregon, April, 1997
- Member, Program Committee, Genome Informatics 1997, Tokyo, Japan, December 1997
- Chair, Program Committee, 2nd Annual International Computational Molecular Biology Conference (RECOMB 98), New York, New York, March, 1998
- Member, Program Committee, 3rd Annual International Computational Molecular Biology Conference (RECOMB 99), Lyon, France, April, 1999
- Member, Program Committee, Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2000), San Francisco, California, January, 2000
- Member, Program Committee, 4th Annual International Computational Molecular Biology Conference (RECOMB 2000), Tokyo, Japan, April, 2000
- Member, Program Committee, The Thirty-Second Annual ACM Symposium on Theory of Computing (STOC 2000), Portland, Oregon, May 21-23, 2000.
- Member, Program Committee, 5th Annual International Computational Molecular Biology Conference (RECOMB 2001), Montreal, Canada, April, 2001
- Co-organizer (with E. Myers and M.S. Waterman), DNA Fragment Assembly, Los Angeles, California, May, 2001
- Member, Program Committee, 6th Annual International Computational Molecular Biology Conference (RECOMB 2002), Washington, DC, April, 2002
- Member, Program Committee, DNA Sequencing and Characterization, Palo Alto, California, May, 2002
- Co-organizer (with S. Tavaré, S. Istrail, and B. Penner), Maps, Sequences, and Genomes. Special Conference devoted to the 60th birthday of Michael Waterman, Los Angeles, California, May, 2002
- Member, Program Committee, 13th Annual Symposium on Combinatorial Pattern Matching, (CPM 2002), Fukuoka, Japan, July 2002

- Member, Program Committee, 1st European Conference on Computational Biology, (ECCB 2002), Saarbrücken, Germany, October 2002
- Member, Program Committee, 7th Annual International Computational Molecular Biology Conference (RECOMB 2003), Berlin, Germany April, 2003
- Member, Program Committee, Intelligent Systems in Molecular Biology (ISMB 2003), Brisbane, Australia, July, 2003
- Member, Organizing Committee, Bertinoro Computational Biology Meeting, Bertinoro, Italy, July 2003
- Member, Program Committee, The Second Asia Pacific Bioinformatics Conference, Dunedin, New Zealand, January, 2004
- Member, Program Committee The Second Asia Pacific Bioinformatics Conference, Dunedin, New Zealand, January, 2004
- Member, Program Committee, 8th Annual International Computational Molecular Biology Conference (RECOMB 2004), San Diego, April, 2004
- Member, Program Committee, RECOMB-Regulatory Genomics workshop, San Diego, California, 2004.
- Member, Program Committee, 8th Annual International Computational Molecular Biology Conference (RECOMB 2004), San Diego, April, 2004
- Member, Program Committee, Eight Annual Conference On Computational Genomics, Boston, Massachusetts, October, 2004
- Member, Program Committee, RECOMB-Comparative Genomics workshop, Bertinoro, Italy, October, 2004
- Member, Program Committee, 9th Annual International Computational Molecular Biology Conference (RECOMB 2005), Boston, May, 2005
- Member, Program Committee, 8th Annual Conference On Computational Genomics, Boston, Massachusetts, November, 2005
- Member, Program Committee, 10th Annual International Computational Molecular Biology Conference (RECOMB 2006), Venice, Italy, April, 2006
- Member, Program Committee, 9th Annual Conference On Computational Genomics, Baltimore, Maryland, October, 2006
- Organizer, Algorithmic Biology, La Jolla, California, November, 2006
- Member, Program Committee, RECOMB-Computational Proteomics Workshop, La Jolla, California, November, 2004
- Member, Program Committee, 11th Annual International Computational Molecular Biology Conference (RECOMB 2007), Oakland, California, April, 2007
- Member, Program Committee, 12th Annual International Computational Molecular Biology Conference (RECOMB 2008), Singapore, April, 2007

Review Panels

- Member, Review board of the Soviet Human Genome Program (Mathematics and Computer Science), 1989-1990
- Member, National Institute of Health Scientific Review Group, Houston, Texas, October 9-11, 1991

- Member, DOE Grant Review Panel, Washington, DC, October 6-9, 1992
- Member, National Institute of Health Scientific Review Group, Washington, DC, January 20-22, 1993
- Member, NSF HPCC Grant Review Panel, Washington, DC, April 26, 1993
- Member, Genome Study Section, National Institutes of Health, Bethesda, Maryland, February 24-26, 1994
- Member, DOE Informatics Panel, Washington, D.C., September, 12-14, 1994
- Member, NSF Theory of Computing Working Group, Las Vegas, Nevada, May 1-2, 1995
- Member, The Swedish Foundation for Strategic Research Advisory Committee, Stockholm, Sweden, September 24-26, 1998
- Member, DIMACS NSF Computational Molecular Biology Activities. Evaluation Committee. October, 1998
- Member, NSF Advisory Board on Computational Biology Activities, Washington, DC, May 1999
- Member, External Review Group, Computer Science Department at Duke University, March, 2003
- External Reviewer, Swedish Foundation for Strategic Research, 2004-2005
- Member, Scientific Advisory Board, Genomics Institute of Singapore, 2005-present
- Member, BDMA Study Section, National Institutes of Health, Bethesda, Maryland, October, 2007

Honors and Awards

- NSF Young Investigator Award, 1994
- Howard Hughes Medical Institute Professor Award, 2006
- UCSD Chancellor Associates Award for Excellence in Research, 2007

Consulting Service

- Hitachi, San Francisco, CA, 1991-1992
- Affymetrix, Santa Clara, CA, 1993-1998
- Millennium Pharmaceuticals, Boston, MA, 1996 - 2000.
- GeneData AG, Basel, Switzerland, 1997-2000 (Scientific Advisory Board)
- Anchorgen, Santa Monica, CA (co-founder), 1998-2004.
- SurroMed, Palo Alto, 2000 - 2001
- X-Mine, Palo Alto, 2000 - 2004 (Scientific Advisory Board)
- Celterra, Carsbad, 2002-2003
- Parity Computing, San Diego, 2004-2005
- Affymetrix, Santa Clara, CA, 2006-2007

Distinguished Lectures at Universities

Transforming Men into Mice

SUNY at Albany, Albany, New York, December 1999

Pattern Discovery in DNA

University of Maryland, College Park, Maryland, November, 2000

Finding Subtle Motifs in DNA sequences

University of California at Irvine, Irvine, California, September, 2002

Transforming Men into Mice

Duke University, Durham, North Carolina, October 2003

Transforming Men into Mice

Carnegie Mellon University, Pittsburgh, Pennsylvania, California, February 2004

Transforming Men into Mice

University of Pennsylvania, Philadelphia, March 2004

Transforming Men into Mice

University of Texas at Dallas, Dallas, Texas, March 2004

Transforming Men into Mice

Harvey-Mudd College, Cleremont, California, September, 2004

Transforming Men into Mice

University of California at Davis, Davis, California, October, 2004

Transforming Men into Mice

University of California at Irvine, Irvine, California, October, 2004

Transforming Men into Mice

Georgia Tech, Atlanta, Georgia, January, 2005

The Third Rebuttal of the Random Breakage Theory

University of Toronto, Toronto, Canada, December, 2006

The Third Rebuttal of the Random Breakage Theory

Brown University, Providence, Rhode Island, December, 2006

The Third Rebuttal of the Random Breakage Theory

University of Maryland, College Park, Maryland, February, 2007

Invited presentations at academic institutions

Convey's equation and paradoxes involving overlapping words

Moscow State University, Moscow, USSR, March 1989

L-tuple sequence reconstruction

Institute for Molecular Genetics, Belgrad, Yugoslavia, November 1989

Algorithms for DNA Sequencing by Hybridization

European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, December 1989.

L-tuple sequence reconstruction

University of Southern California/Department of Mathematics,
Los Angeles, California, June 1990

Multiple sequence alignment with guaranteed error bounds.

University California, Davis/Computer Science Department
Davis, California, December, 1991

SBH and generalized sequencing chips.

Pennsylvania State University, University Park, Pennsylvania, December, 1991

Overlapping words paradox.

San Diego State University/Department of Mathematics
San Diego, California, April, 1992

- Optimal sequencing chips.*
University California Berkeley/Computer Science Department
Berkeley, California, May, 1992
- Conway equation and DNA statistics*
Stanford University/Department of Mathematics
Stanford, California, May, 1992
- Recent advances in Sequencing By Hybridization.*
National Institutes of Health, Bethesda, Maryland, July, 1992
- Towards DNA sequencing chips.*
Pennsylvania State University/Biology Department,
University Park, Pennsylvania, April, 1993
- Towards DNA sequencing chips.*
University of Washington/Department of Molecular Biotechnology
Seattle, Washington, April, 1993
- Genome rearrangements and sorting by reversals.*
University of Southern California/Department of Mathematics
Los Angeles, California, August, 1993
- DNA statistics and the best bet for simpletons*
The Pennsylvania State University/Department of Statistics
University Park, Pennsylvania, September, 1993
- Genome Rearrangements, SBH and double digest problem*
Johns Hopkins University/Computer Science Department
Baltimore, Maryland, October, 1993
- Genome Rearrangements, SBH and double digest problem*
Carnegie Mellon University/Computer Science Department
Pittsburgh, Pennsylvania, October, 1993
- Recent advances in DNA Sequencing by Hybridization*
Baylor College of Medicine, Houston, Texas, November, 1993
- Genome Rearrangements*
Stanford University/Department of Mathematics,
Stanford, California, November, 1993
- Genome Rearrangements, SBH and double digest problem*
Rutgers University/DIMACS
New Brunswick, New Jersey, December, 1993
- DNA statistics and the best bet for simpletons*
Mount Sinai Medical School, New York, New York, December, 1993
- Genome Rearrangements, SBH and double digest problem*
University of Pennsylvania/Computer Science Department
Philadelphia, Pennsylvania, December, 1993
- Genome Rearrangements, SBH and double digest problem*
Polytechnic University/Computer Science Department
Brooklyn, New York, December, 1993
- Towards DNA Sequencing Chips*
Massachusetts Institute of Technology/Whitehead Institute for Biomedical Research
Boston, Massachusetts, February, 1994
- Genome Rearrangements*
Massachusetts Institute of Technology/Computer Science Department
Boston, Massachusetts, February, 1994
- Towards DNA Sequencing Chips*
Boston University/Center for Advanced Biotechnology
Boston, Massachusetts, February, 1994
- Genome Rearrangements*
University of Maryland
College Park, Maryland, April 1994

Towards DNA Sequencing Chips
The Pennsylvania State University
University Park, Pennsylvania, June 1994

Transforming men into mice
University of Southern California
Los Angeles, California, June 1994

Seminar series: 1. Genome rearrangements, 2. DNA chips 3. Multiple alignment
German National Institute for Computer Science (GMD),
St. Augustin, Germany, August, 1994

Seminar series: 1. Towards DNA Sequencing Chips 2. Genomic sequence comparison
Weizmann Institute, Rehovot, Israel, January, 1995

Genome rearrangements
Tel Aviv University, Tel Aviv, Israel, January, 1995

Seminar series: 1. Towards DNA Sequencing Chips 2. Genomic sequence comparison
Washington University, St. Louis, May, 1995

Transforming Mice into Men
Steklov Mathematical Institute, Moscow, Russia, June, 1995

Towards DNA Sequencing Chips
Moscow State University, Moscow, Russia, July, 1995

Genome Rearrangements
INRIA, Paris, France, July, 1995

Genome Rearrangements
Universite de Paris-Sud, Centre d'Orsay, Orsay Cedex, France July, 1995

Towards DNA Sequencing Chips
Ohio State University, Columbus, Ohio, September, 1995

Transforming Mice into Men
Columbia University, New York, New York, September, 1995

A Spliced Alignment Problem: A New Approach to Gene Recognition
Princeton University, Princeton, New Jersey, October, 1995

Transforming Mice into Men
National Institutes of Health, Bethesda, Maryland, December 1995

A New Approach to Gene Recognition
Department of Biology, University of Southern California,
Los Angeles, California, February, 1996

Genome Rearrangements
Department of Computer Science, University of Washington,
Seattle, Washington, March, 1996

Gene Recognition via Spliced alignment
Department of Molecular Biotechnology, University of Washington,
Seattle, Washington, March, 1996

Transforming Mice into Men
Tel Aviv University, Tel Aviv, Israel, May, 1996

New Approaches to Gene Recognition
Pennsylvania State University, State College, PA, November, 1996

Transforming Mice into Men
National Institute for Genetics, Mishima, Japan, December 1996

New Approaches to Gene Recognition
RWCP (Real World Computing Partnership), Tokyo, Japan, December 1996

Genome Rearrangements
Institute for Molecular Biology, Moscow, Russia, June, 1997

Finding genes (children lecture)
Canoga Park High School, Canoga Park, California, March, 1998

Computational proteomics
Institute of Molecular Biology, Moscow, Russia, September 1998

Finding genes and breaking secret codes (children lecture)
Nitanny Valley School, State College, Pennsylvania, December, 1998

Transforming Men into Mice
University of California San Diego, San Diego, California, February 2000

Transforming Men into Mice
University of California at Los Angeles, Los Angeles, California, April, 2000

A New Approach to Sequence Alignment
Institute for Molecular Biology, Moscow, Russia, October, 2000

Assembling Puzzles by Breaking them into Smaller Pieces
Department of Genetics, Washington University, St. Louis, Missouri, May, 2001

Assembling Puzzles by Breaking them into Smaller Pieces
San Diego Supercomputing Center, University of California at San Diego, La Jolla, California, May, 2001

Assembling Puzzles by Breaking them into Smaller Pieces
Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, October, 2001

Assembling Puzzles by Breaking them into Smaller Pieces
Genome Center, Columbia University, New York, New York, October, 2001

Reconstructing Gene Orders in the Ancestral Species
National Cancer Institute, Frederick, Maryland, March, 2002

Finding Subtle Motifs in DNA sequences
University of California at Riverside, Riverside, California, March, 2003

Transforming Men into Mice
King's College, London, England, September, 2003

Transforming Men into Mice
Case Western Reserve University, Cleveland, Ohio, October 2003

Transforming Men into Mice
Institute of Molecular Biology, Moscow, Russia, December, 2003

Transforming Men into Mice
Tel Aviv University, Israel, May 2004

De novo Repeat Classification and Fragment Assembly
University of Hong Kong, Hong Kong, November, 2004

Transforming Men into Mice
City University of Hong Kong, Hong Kong, October, 2004 (colloquium)

Seminar series: 1. Genome rearrangements, 2. Repeat Classification and Fragment Assembly 3. Peptide sequencing
National University of Singapore,
Singapore, November-December, 2004

Genome Rearrangements in Cancer
Genomic Institute of Singapore, Singapore, December, 2004

Genome Rearrangements in Cancer
Georgia State University, Atlanta, Georgia, January, 2005

Transforming Men into Mice
University of Connecticut, Starr, Connecticut, March 2005 (colloquium)

Transforming Men into Mice
Institute of Molecular Biology, Moscow, Russia, June 2005

Blind PTM search via Mass-Spectrometry
National Institute of Health, Bethesda, MD, September, 2005

Blind PTM search via Mass-Spectrometry
National University of Singapore,
Singapore, November, 2005

Fragile versus Random Models of Chromosome Evolution
Genomic Institute of Singapore, Singapore, November, 2005

Proteogenomic Annotations of Bacterial Genomes
Institute of Molecular Biology, Moscow, Russia, November 2006

The Third Rebuttal of the Random Breakage Theory

Computer Center of the Russian Academy of Sciences, Moscow, Russia, November 2006
Comparative Proteogenomics
Institute of Molecular Biology, Moscow, Russia, September 2007

Invited presentations at companies

Multiple filtration in fast database search

Wagner, Inc., Sunnyville, California, April, 1992

Recent advances in DNA sequencing by Hybridization.

Affymetrix, Inc., Santa Clara, California, August, 1993

A Spliced Alignment Problem: A New Approach to Gene Recognition

Sequana, Inc., La Jolla, California, November, 1995

Gene Recognition via Spliced alignment

Amgen, Inc., Thousand Oaks, California, April, 1996

Gene Recognition via Spliced alignment

Millenium Pharmaceutical, Inc., Boston, Massachussets, May, 1996

New Approaches to Gene Recognition

Novartis, Basel, Switzerland, May 1997

Gene hunting without genomic sequencing: the twenty questions game with genes

SmithKline Beecham, King of Prussia, Pennsylvania, August, 1997

Gene hunting without genomic sequencing: the twenty questions game with genes

Millenium Pharmaceuticals, Boston, Massachussets, October, 1997

Transforming Mice into Men

IBM, Yorktown Heights, NY, December 1997

Gene hunting without genomic sequencing: the twenty questions game with genes

Human Genome Sciences, Rockville, Maryland, March 1998

Gene prediction: merging computational and experimental approaches

Ceres, Inc., Malibu, California, September 1998

Finding post-translational modifications in proteins

Sequenom, La Jolla, California, January, 2000

Proteomics in the postgenomic era: finding postranslational modifications in proteins

Celera Genomics, Rockville, Maryland, March, 2000

A New Approach to Sequence Alignment

Paracel, Pasadena, California, July, 2000

Assembling Puzzles by Breaking them into Smaller Pieces

Integrative Genomics, Chicago, Illinois, October, 2000

Finding post-translational modifications in proteins

SurroMed, Palo Alto, California, June, 2001

Finding post-translational modifications in proteins

Syngenta, La Jolla, California, October, 2001

Assembling Puzzles by Breaking them into Smaller Pieces

Omnigon, Carsbad, California, November, 2001

Human and mouse genomic sequences reveal evidence against random breakage model of chromosome evolution

Celera Genomics, Rockville, MD, November, 2003

Human and mouse genomic sequences reveal evidence against random breakage model of chromosome evolution

The Center for Genomic Research, Rockville, MD, November, 2003

Computational Mass-Spectrometry

Strand Genomics, Bangalore, India, December, 2004

Short Reads Fragment Assembly

Helicos, Cambridge, Massachusetts, June, 2006

Short Reads Fragment Assembly

Complete Genomics, Sunnyvale, California, September, 2006

Keynote presentations at professional meetings

Towards DNA sequencing chips

19th International symposium on Mathematical Foundations in Computer Science
Kosice, Slovakia, August, 1994 (keynote)

Open Problems in Computational Molecular Biology

3rd Israel Symposium on Theory of Computing and Systems
Tel Aviv, January, 1995 (keynote)

Transforming Mice into Men

Workshop on String Processing
Recife, Brazil, August, 1996 (keynote)

Transforming Mice into Men

Genome Informatics 1996, Tokyo, Japan, December 1996 (keynote)

Transforming Mice into Men

American Mathematical Society meeting,
Corvallis, Oregon, April, 1997 (invited address)

Gene hunting without genomic sequencing: the twenty questions game with genes

German Conference on Bioinformatics,
Kloster Irsee, Germany, September, 1997 (keynote)

Transforming Mice into Men

DIMACS 10th Anniversary Celebration, Piscataway, New Jersey, October, 1998
(keynote)

De novo protein sequencing

Symposium on String Processing and Information Retrieval (SPIRE 99),
Cancun, Mexico, September 99
(keynote)

Transforming Men into Mice

Canadian Mathematical Society Winter Meeting, Montreal, Canada, December 99
(keynote)

Assembling Puzzles by Breaking them into Smaller Pieces

SIAM Conference on Discrete Mathematics, San Diego, CA, August, 2003
(keynote)

Transforming Men into Mice: Nadeau-Taylor chromosomal breakage theory revisited

New Frontiers In Biological Sciences Symposium, Los Angeles, CA, April, 2003
(keynote)

Transforming Men into Mice

24th Conference on Foundations of Software Technology and Theoretical Computer Science,
Chennai, India, December 2004 (keynote)

Transforming Men into Mice: Nadeau-Taylor chromosomal breakage theory revisited

Annual Meeting of American Mathematical Society, Atlanta, Georgia, January 2005
(invited address)

Fragile Breakage Model versus Random Breakage Model

Intelligent Systems in Molecular Biology (ISMB 2005) , Detroit, Michigan, June 2005
(keynote)

Spectral Networks

Asia Pacific Bioinformatics Conference (APBC 2007), Hong Kong, China, January 2007
(keynote)

Fragile Breakage Model versus Random Breakage Model

German Conference on Bioinformatics,

Potsdam, Germany, September 2007 (keynote)
Shotgun Protein Sequencing
Biomedical Computation at Stanford,
Stanford, California, October 2007 (keynote)

Invited presentations at professional meetings

Combinatorial methods for SBH

Sequencing by Hybridization, Washington,DC, May 1991

Nucleotide sequences versus Markov chains.

Open Problems in Computational Molecular Biology 1991,
Telluride, Colorado, June 1991

Multiple alignment and communication cost

3rd Annual Symposium 'Combinatorial Pattern Matching 1992',
Tucson, Arizona, April,1992

Generalized sequence alignment.

3rd Annual Symposium 'Combinatorial Pattern Matching 1992',
Tucson, Arizona, April,1992

Sequencing by hybridization

Computational biology - cutting edge'
Irvine, California, May, 1992

Overlapping word paradox and DNA statistics.

Supercomputing and Complex Genome Analysis 1992,
St. Petersburg Florida, June, 1992

Multiple sequence alignment with guaranteed error bounds.

Algorithms for DNA Sequence Comparison,
Albuquerque, New Mexico, November 1992

Nucleotide sequences versus Markov models.

Mathematics and Molecular Biology III. Santa Fe, New Mexico, November 1992

A fast filtration algorithm for the substring matching problem.

4th Annual Symposium "Combinatorial Pattern Matching 93",
Padova, Italy, June,1993

Multiple sequence comparison and n-dimensional image recognition.

4th Annual Symposium "Combinatorial Pattern Matching 93",
Padova, Italy, June,1993

Towards DNA sequencing chips.

Open Problems in Computational Molecular Biology 1993
Telluride, Colorado, July, 1993

DNA inhomogeneity and overlapping words paradox

DNA sequence analysis, Stanford, California, August, 1993

How to combine SBH with additional biochemical experiments

Second International SBH Workshop, Houston, Texas, October, 1993

Sorting by reversals

Combinatorial Methods for Genome Rearrangements,
Los Angeles, California, March 1994

Towards DNA sequencing chips

1st World Congress on Computational Medicine, Public Health and Biotechnology,
Austin, Texas, April 1994

Towards computational theory of genome rearrangements

Bioinformatics and Complex Genome Analysis 1994,
Tallahassee, Florida, June, 1994

Genome rearrangements
Computational Molecular Biology, Albuquerque, New Mexico, June, 1994

Rearrangement of maps, sequences and genomes
IMA Summer Program in Molecular Biology, University of Minnesota,
Minneapolis, Minnesota, July, 1994

Rearrangement of maps, sequences and genomes
Combinatorial Methods in DNA Mapping and Sequencing, DIMACS, Rutgers University
New Brunswick, New Jersey, October, 1994

Towards DNA sequencing chips
Combinatorial Structures in Molecular Biology, DIMACS, Rutgers University
New Brunswick, New Jersey, November, 1994

Genomic sequence comparison
DNA Sequence Alignment,
Princeton, New Jersey, November, 1994

Transforming Mice into Men
Evolutionary Molecular Biology, Canadian Research Council
Montreal, Canada, August, 1995

Towards DNA chips
Microfabrication Technology, Cambridge Healthtech Institute
San Francisco, California, September, 1995

Genome Rearrangements
4th DIMACS Implementation Challenge, DIMACS, Rutgers University
New Brunswick, New Jersey, September, 1995

A Spliced Alignment Problem: A New Approach to Gene Recognition
DIMACS Gene Recognition Workshop
Philadelphia, Pennsylvania, October, 1995

Genome Rearrangements
Mathematics and Molecular Biology IV
Santa Fe, New Mexico, November, 1995

Gene Recognition: Combinatorics versus Statistics
2nd Sandia Workshop on Computational Molecular Biology
Albuquerque, New Mexico, March, 1996

Genome Rearrangements
Conference on Computational Molecular Biology to honor 50th anniversary of ENIAC
Princeton, New Jersey, May, 1996

Gene Recognition
Israeli Computational Molecular Biology Conference
Nasholim, Israel, June, 1996

Las Vegas Algorithms for Gene Recognition
1st International Conference on Computational Molecular Biology,
Santa Fe, New Mexico, January, 1997

Transforming Mice into Men
Computational Biology Workshop, Zurich, Switzerland, May, 1997

Gene hunting without genomic sequencing: the twenty questions game with genes
Gene prediction in silico, Atlanta, GA November, 1997

Gene hunting without genomic sequencing: the twenty questions game with genes
Functional Gene Analysis, Cambridge Healthtech Institute,
San Francisco, California, February, 1998

Gene hunting without genomic sequencing: the twenty questions game with genes
Understanding the Genome: technological and mathematical aspects,
Berkeley, California, May 1998

Computational challenges in gene hunting
Computational Biology School, Udine, Italy, June 1998

De novo protein sequencing by mass-spectrometry

- The 46th ASMS Conference on Mass-Spectrometry, Orlando, Florida, June 1998
- De novo protein sequencing by mass-spectrometry*
 Computational Genomics, The Institute for Genomic Research,
 Rockville, Maryland, November, 1998
- Genome rearrangements*
 Lipari Computational Biology School, Lipari, Italy. June, 1999
- Finding Post-Translational Protein Modifications*
 Bioinformatics 2000, Elsinor, Denmark, May, 2000
- Finding Post-Translational Protein Modifications*
 Cambridge Healthtech Institute Conference on Bioinformatics, San Francisco, California, June, 2000
- Finding Post-Translational Protein Modifications*
 Beyond the Genome, Berkeley, California, June, 2000
- Edgar Allan Poe and Computational Molecular Biology*
 Genome Research Perspectives for Helmholtz Association, Munich, Germany, November, 2000
- Assembling Puzzles by Breaking them into Smaller Pieces*
 Genome Sequencing and Biology, Cold Spring Harbor, New York, May, 2001
- Assembling Puzzles by Breaking them into Smaller Pieces*
 Beyond Genome, Cambridge Healthtech Institute Conference, San Francisco, California, June, 2001
- Assembling Puzzles by Breaking them into Smaller Pieces*
 Post-genomic Bioinformatics, Madrid, Spain, April, 2002
- Assembling Puzzles by Breaking them into Smaller Pieces*
 Maps, Sequences, and Genomes, Los Angeles, CA, May, 2002
- Computational mass-spectrometry*
 Keystone Symposium: "Proteomics: Technology and Applications" Keystone, Colorado, March 2003.
- Short Read Assembly*
 RECOMB-Genomic Technologies Satellite Meeting, Stanford, California, May 2004.
- Transforming Men into Mice*
 Human Genome Organization (HUGO) Meeting, Singapore, November 2004.

Publications

- Pevzner P.A., M. Yu. Borodovsky and A.A. Mironov (1989) Linguistics of Nucleotide Sequences I: the significance of deviations from mean statistical characteristics and prediction of the frequencies of occurrence of words. *J. Biom. Str. & Dyn.*, **6**, 1013-1026
- Pevzner P.A., M. Yu. Borodovsky and A.A. Mironov (1989) Linguistics of Nucleotide Sequences II: stationary words in genetic texts and the zonal structure of DNA. *J. Biom. Str. & Dyn.*, **6**, 1027-1040
- Pevzner P.A., V.P. Veiko (1989) Cycles in graphs and oligonucleotide linkers. *Molek. Biol.*, **4**, 1075-1079
- Pevzner P.A. (1989) *L*-tuple DNA sequencing: computer analysis. *J. Biomol. Str. & Dyn.*, **7**, 63-74
- Borodovsky M.A., Pevzner P.A. (1990) Statistical analysis of genetics texts. *Computer Analysis of Genetics Texts*, Chapter 2, Ed. M.D. Frank-Kamenetzki, Nauka, Moscow, 36-80 (In Russian)
- Pevzner P.A. (1990) DNA physical mapping. *Computer Analysis of Genetic Texts*, Chapter 5, Ed. M.D. Frank-Kamenetzki, Nauka, Moscow, 154-188 (In Russian)
- Pevzner P.A. (1990) Methods of filtration and database search. *Biopolymers & Cell*, **6**,7-13.
- Kozhukhin C.G.,Pevzner P.A. (1991) Genome inhomogeneity is determined mainly by WW and SS dinucleotides. *Comp. Appl. in Biosci.*, **7**, 39-49.

- Pevzner P.A., Y. Lysov, K. Khrapko, V., Belyavsky A.V., Florentév and A. Mirzabekov (1991) Improved matrices for Sequencing by Hybridization., *J. Biomol. Struct. and Dynamics*, **9**, 399-410.
- M.Gelfand, C.Kozhukhin, Pevzner P.A. (1992) Extendable words in nucleotide sequences *Comp. Appl. in Biosci.*, **8**, 129-135
- Pevzner P.A. (1992) Statistical distance between texts and filtration methods in rapid similarity search algorithm. *Comp. Appl. in Biosci.*,**8**,121-127
- Pevzner P.A. (1992) Nucleotide sequences versus Markov models. *Computers and Chemistry*, **16**, 103-106,
- Pevzner P.A. (1992) DNA physical mapping, flows in networks and minimum cycles mean in graphs. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, **8**, 99-112
- Pevzner P.A. (1992) Multiple alignment, communication cost, and graph matchings. *SIAM J. Appl. Math.*, **52**, 1763-1779
- P.A. Pevzner, M.Waterman (1992) Generalized longest common subsequence problem and Hilbert basis. A.Apostolico, M.Crochermore, Z.Galil, U.Manber (eds.) *Combinatorial Pattern Matching 1992. Lecture Notes in Computer Science*, **644**, 79-89
- Pevzner P.A. (1992) Multiple sequence alignment and communication cost. A. Apostolico, M. Crochermore, Z. Galil, U. Manber (eds.) *Combinatorial Pattern Matching 1992. Lecture Notes in Computer Science*, **644**, Springer-Verlag, 1992, 205-213
- Pevzner P.A., M.Waterman (1993) Generalized sequence alignment and duality. *Adv. in Appl. Math.*, **14**, 139-171.
- Pevzner P.A., M.Waterman (1993) A fast filtration algorithm for the substring matching problem. A.Apostolico, M.Crochermore, Z.Galil, U.Manber (eds.) *Combinatorial Pattern Matching 1993. Lecture Notes in Computer Science*, **684**, 197-214 (appeared in *Algorithmica*, **13** (1995), 135-154)
- M.Vingron, Pevzner P.A. (1993) Multiple sequence alignment and n -dimensional image reconstruction. A.Apostolico, M.Crochermore, Z.Galil, U.Manber (eds.) *Combinatorial Pattern Matching 1993. Lect. Notes in Comp. Sci.*, **684**, 243-253 (appeared in *Advances in Applied Mathematics*, **16** (1995), 1-22)
- Pevzner P.A. (1993) Overlapping word paradox and Conway equation. In C. Cantor, J. Fickett, R. Robbins and H. Lim (eds.) *Supercomputing, Bioinformatics and Complex Genome Analysis*, World Scientific, 71-78
- M. Gelfand, P.A. Pevzner (1993) Editor's Foreword. *Bio Systems*, **30**, 1-4
- Pevzner P.A. (1994) 3-non-crossed families and multicommodity flows. In A.K. Kelmans (ed.) *Selected Topics in Discrete Mathematics*, **158**, American Mathematical Society, 201-206
- Pevzner P.A. (1994) Packing branchings in weighted graphs. In A.K. Kelmans (ed.) *Selected Topics in Discrete Mathematics*, **158**, American Mathematical Society, 1994, 185-200
- Feldman W., Pevzner P.A. (1994) Gray Code Masks for DNA Sequencing by Hybridization *Genomics*, **23**, 233-235
- Pevzner P.A. (1994) Combinatorial methods for DNA physical mapping. *IEEE Computers in Medicine and Biology* **13**, 146-150
- V.Bafna, E.Lawler, Pevzner P.A. (1994) Approximation algorithms for multiple alignment. M. Crochermore and D.Gusfield (eds.) *Combinatorial Pattern Matching 1994, Lecture Notes in Computer Science*, **807**, 43-53
- X.Huang, Pevzner P.A., W.Miller (1994) Parametric recomputing in alignment graphs. M. Crocher-

- more and D.Gusfield (eds.) *Combinatorial Pattern Matching 1994, Lecture Notes in Computer Science*, **807**, 87-101
- Pevzner P.A., Lipshutz R. (1994) Towards DNA Sequencing Chips. *19th Symposium on Mathematical Foundations of Computer Science, Kosice, Slovakia, Lecture Notes in Computer Science*, **841**, 143-158
- Pevzner P.A. (1994) Rearrangements of DNA sequences and SBH. *Computers and Chemistry*, **18**, 221 - 223
- Pevzner P.A. (1995) DNA physical mapping and alternating Eulerian paths in colored graphs. *Algorithmica*, **13**, 77-105
- Pevzner P.A., Waterman M.S. (1995) Open combinatorial problems in computational molecular biology. *Proc. of the Third Israel Symposium on Theory of Computing and Systems, January, 1995, Tel Aviv, Israel*, IEEE Computer Society Press, 158-173
- Pevzner P.A. (1995) Editor's Foreword. *J. of Computational Biology*, **2**, 153-158
- Hannenhalli S., C. Chappay, Koonin E., Pevzner P.A. (1995) Genome rearrangements and evolution of herpes viruses. *Proc. of 3rd International Conference on Bioinformatics and Complex Genome Analysis*, World Scientific, 91-106
- Hannenhalli S., C. Chappay, Koonin E., Pevzner P.A. (1995) Genome sequence comparison and scenarios for gene rearrangements: a test case. *Genomics*, **30**, 299-311
- Mironov A.A., Alexandrov N.N., Bogodarova N.Y., Grigoriev A., Lebedev V., Lunovskaya-Gurova, Grigoriev A., Lebedev V.Truchan M., Pevzner P.A. (1995) DNASUN: a software package for the biotechnology laboratory, *Computer Applications in Biological Sciences*, **11**, 331-335
- Bafna V. and Pevzner P.A. (1995) Sorting by reversals: genome rearrangements in plant organelles and evolutionary history of X chromosome. *Molecular Biology and Evolution*, **12**, 239-246
- Hannenhalli S., Pevzner P.A. (1995) Transforming cabbage into turnip (polynomial algorithm for sorting signed permutations by reversals). *Proc. of the 27th Annual Symposium on the Theory of Computing (STOC 95)*, 178-189 Las Vegas, Nevada
- Hannenhalli S., Pevzner P.A. (1995) Transforming mice into men (polynomial algorithm for genomic distance problem). *Proc. of the 36 Annual Symposium on Foundations of Computer Science (FOCS 95)*, Milwaukee, Wisconsin, 581-592
- Hannenhalli S., Pevzner P.A. (1995) Towards Computational Theory of Genome Rearrangements. In J. van Leeuwen, (ed.) *Computer Science Today: Recent Trends and Developments*. Lecture Notes in Computer Science, **1000**, 184-202 (invited paper in a jubilee volume)
- Hannenhalli S., Pevzner P.A. (1996) To cut... or not to cut: applications of comparative physical maps in molecular evolution. *Proc. of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 96)*, Atlanta, Georgia, 304-313
- V.Bafna, Pevzner P.A. (1996) Genome rearrangements and sorting by reversals. *SIAM J. Computing*, **25**, 272-289. (preliminary version appeared in Proc. of 34th IEEE Symposium on Foundations of Computer Science (FOCS 93) , November 3-5, 1993, Palo Alto, CA, 148-157)
- Hannenhalli S., W. Feldman, H. Lewis, S.Skienna, Pevzner P.A. (1996) Positional Sequencing by Hybridization. *Computer Appl. in Biol. Sciences* **12**, 19-24
- M. Gelfand, A. Mironov, P.A. Pevzner (1996) Spliced Alignment Problem: A New Approach to Gene Recognition. *Combinatorial Pattern Matching 1996, Lecture Notes in Computer Science*, **1075**, 141-158

- M.Gelfand, A. Mironov, P.Pevzner (1996) Gene Recognition via Spliced Alignment *Proc. National Academy of Sciences USA*, **93**, 9061-9066
- S.H. Sze, P.A. Pevzner (1997) Las Vegas Algorithms for Gene Recognition. *Proc. of the 1st Annual International Conference on Computational Molecular Biology*, Santa Fe, New Mexico, January, 1997, 300-309
- Belyi I., Pevzner P.A. (1997) Software for DNA sequencing by hybridization. *Computer Applications in Biological Sciences*, **13**, 205-210
- A.V. Aho, D.S. Johnson, R.M. Karp, S.R. Kosaraju, C.C. McGeoch, C.H. Papadimitriou, P.A. Pevzner (1997) Emerging Opportunities for Theoretical Computer Science, *SIGACT News*, **28**, 65-74
- V.Bafna, E.Lawler, Pevzner P.A. (1997) Approximation algorithms for multiple alignment. *Theoretical Computer Science*, **182**, 233-244
- S.H. Sze, P.A. Pevzner (1997) Las Vegas Algorithms for Gene Recognition: Suboptimal and Error-Tolerant Spliced Alignment. *J. Comp. Biol.*, **4**, 297-310
- G.Xu, S.H. Sze, C.P.Lin, P.A. Pevzner, N. Arnheim. (1998) Gene Hunting without Sequencing Genomic Clones: Finding Exon Boundaries in cDNAs. *Genomics*, **47**, 171-179
- Sze S.-H., Roytberg M.A., Gelfand M.S., Mironov A.A., Astakhova T.V. and Pevzner P.A. (1998) Algorithms and software for support of gene identification experiments. *Bioinformatics*, **14**, 14-19
- A. Mironov, M. Roytberg, P.A. Pevzner, M.Gelfand, (1998) Performance Guarantee Algorithms for Gene Recognition. *Genomics* **51**, 332-339
- Bafna V., Pevzner P.A. (1998) Sorting by transpositions. *SIAM J. Discrete Math.*, **11**, 224-240 (preliminary version appeared in *Proceedings of the Sixth Annual Symposium on Discrete Algorithms (SODA 95)*, San Francisco, CA, 614-623)
- L. A. Goldberg, P. W. Goldberg, M. S. Paterson, P. A. Pevzner, S. Cenk Sahinalp, and E. Sweedyk: (1999) Complexity of Gene Placement. *Tenth Annual ACM-SIAM Symposium on Discrete Algorithms*, January 17-19, 1999 Baltimore, Maryland.
- V.Dancik, T.Addona, K.Clauser, J.Vath, P.Pevzner (1999) *De novo* protein sequencing via tandem mass-spectrometry. *J. Comp. Biol.*, **6**, 327-341
- E. Hubbell, P. A. Pevzner (1999) Fidelity probes for oligonucleotide arrays. Seventh International Conference on Intelligent Systems for Molecular Biology (ISMB 99), Heidelberg, Germany, August, 1999, 113-116
- A. Mironov, P.A. Pevzner (1999) SST versus EST in gene recognition. *Microbial and Comparative Genomics*, **4**, 254-260
- S. Hannenhalli, P.A. Pevzner (1999) Transforming cabbage into turnip: a polynomial algorithm for sorting signed permutations by reversals. *Journal of ACM*, **46**, 1-27
- P.A. Pevzner, V.Dancik, C. Tang (2000) Mutation- and modification-tolerant protein identification via tandem mass-spectrometry. *J. Comp. Biology*, **7**, 777-787 (conference version appeared in the Fourth International Conference on Computational Molecular Biology RECOMB 2000, Tokyo, Japan, April 2000, 231-236)
- V. Bafna, D. Beaver, M. Furer, P.A. Pevzner. Circular permutations and genome shuffling. In *Comparative Genomics - Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and the Evolution of Gene Families*. (eds. D. Sankoff and J. Nadeau), Kluwer, 2000.
- P. A. Pevzner, S. H. Sze (2000) Combinatorial algorithm for finding subtle signals in DNA sequences.

- Eighth International Conference on Intelligent Systems for Molecular Biology (ISMB 2000), San Diego, California, August 2000, 269-278
- P.A. Pevzner, H. Tang (2001) Fragment assembly with double-barreled data. *Bioinformatics*, (Proceedings of ISMB 2001) Suppl 1:S225-33.
- A.N. Arslan, O. Egecioglu, P.A. Pevzner. (2001) A new approach to sequence comparison: normalized sequence alignment. *Bioinformatics* **17**, 327-337 (conference version appeared in the Fifth International Conference on Computational Molecular Biology RECOMB 2001, Montreal, Canada, April 2001)
- P.A. Pevzner, Z. Mulyukov, V. Dancik, C.L. Tang (2001) Efficiency of database search for identification of mutated and modified proteins via mass spectrometry. *Genome Res.* **11**, 290-299.
- P.A. Pevzner (2001) The best bet for simpletons *Mathematical World* **17**, American Mathematical Society, 123-131
- L. A. Goldberg, P. W. Goldberg, M. Paterson, P. Pevzner, S. C. Sahinalp, E. Sweedyk. (2001) The Complexity of Gene Placement *Journal of Algorithms*, **41**, 225-243
- P.A. Pevzner (2001) Assembling puzzles from preassembled pieces. *Genome Res.*, **11**, 1461-1462
- P.A. Pevzner, H. Tang, M.S. Waterman (2001) An Eulerian path approach to DNA fragment assembly *Proceedings of the National Academy of Sciences*, **98**, 9748-9753 (extended abstract appeared in the Fifth International Conference on Computational Molecular Biology RECOMB 2001, Montreal, Canada, April 2001)
- Z. Mulyukov and P.A. Pevzner (2002) EULER-PCR: finishing experiments for repeat resolution in DNA sequencing. Pacific Symposium on Biocomputing (PSB 2002), Kauai, Hawaii, 199-210
- S.H. Sze, M.S. Gelfand, and P.A. Pevzner (2002) Finding subtle motifs in DNA sequences. Pacific Symposium on Biocomputing (PSB 2002), Kauai, Hawaii, 235-246
- G. Bourque and P.A. Pevzner (2002) Genome-scale evolution: reconstructing gene orders in the ancestral species. *Genome Research*, **12**, 26-36.
- S. Heber, M. Alekseyev, S.H. Sze, H. Tang and P.A. Pevzner Splicing Graphs and EST Assembly Problem. (2002) *Bioinformatics*, **18**, S181-188 (Proceedings of ISMB 2002)
- E. Eskin and P.A. Pevzner. (2002) Finding composite regulatory patterns in DNA sequences. *Bioinformatics*, **18**, S354-363 (Proceedings of ISMB 2002)
- U. Keich and Pavel A. Pevzner. (2002) Finding motifs in the twilight zone. *Bioinformatics*, **18**, 1374-1381 (extended abstract appeared in the Sixth International Conference on Computational Molecular Biology RECOMB 2002, Washington, DC, April 2002)
- U. Keich and Pavel A. Pevzner. (2002) Subtle motifs: defining the limits of motif finding algorithms. *Bioinformatics*, **18**, 1382-1390
- S. Hannenhalli, E. Hubbell, R. Lipshutz, P.A. Pevzner (2002) Combinatorial Algorithms for Design of DNA Arrays. In *Chip Technology, Adv. Biochem. Eng. Biotechnol.* **77**, 1-19
- A. B. Kahng, I. I. Mandoiu, P. A. Pevzner, S. Reda, A. Z. Zelikovsky. (2002) Border Length Minimization in DNA Array Design. *Workshop on Algorithms in Bioinformatics (WABI 2002)*, Lecture Notes in Computer Science Series **2452**, 435-448
- V. Lunyak, R. Burgess, G. Prefontaine, C. Nelson, S.H. Sze, J. Chenoweth, P. Schwartz, P.A. Pevzner, C. Glass, G. Mandel, and M. G. Rosenfeld. (2002) Corepressor depending silencing of chromosomal regions encoding neuronal genes. *Science*, **298**, 1747-1752
- R. Waterston, ..., P.A. Pevzner. G. Tesler, ..., E. Lander. (2002) Initial sequencing and analysis of

the mouse genome. *Nature*, **420**, 520-562

E. Eskin, U. Keich, M.S. Gelfand, P.A. Pevzner. (2003) Genome-wide analysis of bacterial promoter regions. Pacific Symp. Biocomput, Hawaii, 29-40.

P.A. Pevzner and G. Tesler (2003) Genome rearrangements in mammalian genomes: Lessons from human and mouse genomic sequences. *Genome Research*, **13**, 37-45

P.A. Pevzner and G. Tesler (2003) Transforming men into mice: the Nadeau-Taylor chromosomal breakage model revisited. Seventh International Conference on Computational Molecular Biology RECOMB 2003, Berlin, Germany, April 2003

A. B. Kahng, I. I. Mandoiu, P. A. Pevzner, S. Reda and A. Z. Zelikovsky. Engineering a scalable placement heuristic for DNA probe arrays. Seventh International Conference on Computational Molecular Biology RECOMB 2003, Berlin, Germany, April 2003

P.A. Pevzner and G. Tesler (2003) Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proc Natl Acad Sci U S A*. **100**, 7672-7677

A. Price, S. Ramabhadran, P.A. Pevzner. (2003) Finding Subtle Motifs by branching from sample strings. *Bioinformatics*, **19**, SS149-155

B. Raphael, S. Volik, C. Collins, and P.A. Pevzner (2003) Reconstructing Tumor Genome Architectures *Bioinformatics*, **19**, SS167-SS171

W.J. Murphy, G. Bourque, G. Tesler, P. Pevzner., S.J. O'Brien. (2003) Reconstructing the genomic architecture of mammalian ancestors using multispecies comparative maps. *Human Genomics*, **1**, 30-40

G. Andelfinger, C. Hitte, L. Etter, R. Guyon, G. Bourque, G. Tesler, P. Pevzner, E. Kirkness, F. Galibert, D.W. Benson (2004) Detailed Four-Way Comparative Mapping and Gene Order Analysis of the Canine *ctvm* Locus Reveals Fragile Sites Driving Chromosome Rearrangements. *Genomics*, **83**, 1053-1062.

G. Bourque, P. Pevzner, G. Tesler. (2004) Reconstructing the Genomic Architecture of Ancestral Mammals: Lessons from Human, Mouse, and Rat Genomes. *Genome Research*, **14**, 507-516

Gibbs,..., P.A. Pevzner,..., F. Collins (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493-521.

P.A. Pevzner, H. Tang and G. Tesler (2004) De novo repeat classification and fragment assembly. Eight International Conference on Computational Molecular Biology RECOMB 2004, San Diego, March 2004

P.A. Pevzner, H. Tang. G. Tesler (2004) De novo repeat classification and fragment assembly. *Genome Res.* **14**, 1786-1796

J. Leipzig, P.A. Pevzner, S. Heber. (2004) The Alternative Splicing Gallery (ASG): bridging the gap between genome and transcriptome. *Nucleic Acids Res.* **32**, 3977-3983.

A.B. Kahng, I.I. Mandoiu, P.A. Pevzner, S. Reda, A.Z. Zelikovsky (2004) Scalable heuristics for design of DNA probe arrays. *J Comput Biol.*, **11**, 429-447

B.J. Raphael, P.A. Pevzner. (2004) Reconstructing tumor amplicomes. *Bioinformatics* (ISMB 2004 Special Issue), **20**, Suppl 1:I265-I273

P.A. Pevzner (2004) Educating biologists in the 21st century: bioinformatics scientists vs. bioinformatics technicians. *Bioinformatics*, **20**, 2159-2161

M. Chaisson, P.A. Pevzner, H. Tang H. (2004) Fragment assembly with short reads. *Bioinformatics*, **20**, 2067-2074

- N. Bandeira, H. Tang, V. Bafna, P.A. Pevzner P. (2004) Shotgun protein sequencing by tandem mass spectra assembly. *Anal Chem.*, **76**, 7221-33
- L.W. Hillier, ..., P.A. Pevzner, ..., Wilson R.K. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695-716.
- W.J. Murphy, P.A. Pevzner, S.J. O'Brien (2004) Mammalian phylogenomics comes of age. *Trends Genet.*, **20**, 631-9.
- B. Raphael, D. Zhi, H. Tang, P.A. Pevzner (2004) A novel method for multiple alignment of sequences with repeated and shuffled elements. *Genome Res.* **14**, 2336-46.
- A. L. Price, E. Eskin, P.A. Pevzner (2004) Whole-genome analysis of Alu repeat elements reveals complex evolutionary history. *Genome Res.*, **14**, 2245-52.
- M. Alekseyev and P.A. Pevzner. (2004) Genome Halving Problem revisited. *Lecture Notes in Computer Science*, **3328**, 1-15
- G. Bourque, E.M. Zdobnov, P. Bork, P.A. Pevzner, G. Tesler (2005) Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. *Genome Res.*, **15**, 98-110.
- A. Frank and P.A. Pevzner (2005) PepNovo: De Novo Peptide Sequencing via Probabilistic Network Modeling. *Analytical Chemistry*, **77**, 964 - 973
- A. Price, N.C. Jones, and P.A. Pevzner (2005) De Novo Identification of Repeat Families in Large Genomes. *Bioinformatics* **21**, Suppl 1:i351-i358 (ISMB special issue).
- A. Frank, S. Tanner, V. Bafna, P.A. Pevzner (2005) Peptide sequence tags for fast database search in mass-spectrometry. *J. of Proteome Research*, **4**, 1287-95.
- S. Tanner, H. Shu, A. Frank, M. Mumby, P.A. Pevzner, V. Bafna (2005) InsPecT: Fast and accurate identification of post-translationally modified peptides from tandem mass spectra, *Analytical Chemistry*, **77**, 4626-4639.
- W. J. Murphy, D. M. Larkin, A. Everts-van der Wind, G. Bourque, G. Tesler, L. Auvil, J. E. Beever, B. P. Chowdhary, F. Galibert, L. Gatzke, C. Hitte, D. Milan, S. N. Meyers, E. A. Ostrander, G. Pape, H. G. Parker, T. Raudsepp, M. B. Rogatcheva, L. B. Schook, L. C. Skow, M. Welge, J. E. Womack, S. J. O'Brien, P. A. Pevzner, H.A. Lewin (2005) Dynamics of Mammalian Chromosome Evolution Inferred from Multispecies Comparative Maps. *Science*, **309**, 613-617.
- D. Tsur, S. Tanner, E. Zandi, V. Bafna, P.A. Pevzner. (2005) Identification of post-translational modifications by blind search of mass spectra. *Nature Biotechnology*, **23**, 1562-7.
- G. Bourque, G. Tesler, P.A. Pevzner (2006) The convergence of cytogenetics and rearrangement-based models for ancestral genome reconstruction. *Genome Res.*, **16**, 311-313
- D. Zhi, B.J. Raphael, A.L. Price, H. Tang, P.A. Pevzner (2006) Identifying repeat domains in large genomes. *Genome Biol.*, **7**, R7.
- Q. Peng, P.A. Pevzner, G. Tesler (2006) The fragile breakage versus random breakage models of chromosome evolution. *PLoS Comput Biol.* **2**, e14.
- S. Tanner, P.A. Pevzner, V. Bafna (2006) Unrestrictive identification of post-translational modifications through peptide mass spectrometry. *Nat Protoc.* **1**, 67-72.
- P.A. Wilmarth, S. Tanner, S. Dasari, S.R. Nagalla, M.A. Riviere, V. Bafna, P.A. Pevzner, L.L. David (2006) Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? *J Proteome Res.*, **5**, 2554-66.

- N. Wielsch, H. Thomas, V. Surendranath, P. Waridel, A. Frank, P. Pevzner, A. Shevchenko (2006) Rapid validation of protein identifications with the borderline statistical confidence via de novo sequencing and MS BLAST searches *J Proteome Res.*, **5**, 2448-56.
- N.C. Jones, P.A. Pevzner (2006) Comparative genomics reveals unusually long motifs in mammalian genomes. *Bioinformatics*, **22**, e236-42 (ISMB special issue)
- D. Zhi, S.S. Krishna, H. Cao, P. Pevzner, A. Godzik A. (2006) Representing and comparing protein structures as paths in three-dimensional space. *BMC Bioinformatics* **7** 460
- S. Volik, B.J. Raphael, G. Huang, M.R. Stratton, G. Bignel, J. Murnane, J.H. Brebner, K. Bajsarowicz, P.L. Paris, Q. Tao, D. Kowbel, A. Lapuk, D.A. Shagin, I.A. Shagina, J.W. Gray, J.F. Cheng, P.J. de Jong, P. Pevzner, C. Collins (2006) Decoding the fine-scale structure of a breast cancer genome and transcriptome. *Genome Res.*, **16**, 394-404
- K.F. Chong, K. Ning, H.W. Leong, P. Pevzner (2006) Modeling and characterization of multi-charge mass spectra for Peptide sequencing. *J Bioinform Comput Biol.*, **4**, 1329-52
- M.J. Chaisson, B.J. Raphael, P.A. Pevzner (2006) Microinversions in mammalian evolution. *Proc Natl Acad Sci U S A*, **103**, 19824-9.
- M.A. Alekseyev and P.A. Pevzner (2007) Colored de Bruijn graphs and the genome halving problem. *IEEE/ACM Trans Comput Biol Bioinform* **4**, 98-107
- D. Zhi, U. Keich, P. Pevzner, S. Heber, H. Tang (2007) Correcting base-assignment errors in repeat regions of shotgun assembly. *IEEE/ACM Trans Comput Biol Bioinform.*, **4**, 54-64.
- A.M. Frank, M.M. Savitski, M.L. Nielsen, R.A. Zubarev, P.A. Pevzner (2007) De novo peptide sequencing and identification with precision mass spectrometry. *J Proteome Res.*, **6**, 114-23.
- N. Gupta, S. Tanner, M. Jaitly, J.N. Adkins, M. Lipton, R. Edwards, M. Romine, A. Osterman, V. Bafna, R.D. Smith, P.A. Pevzner (2007) Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res.*, **17**, 1362-77.
- P. Waridel, A. Frank, H. Thomas, V. Surendranath, S. Sunyaev, P. Pevzner, A. Shevchenko (2007) Sequence similarity-driven proteomics in organisms with unknown genomes by LC-MS/MS and automated de novo sequencing. *Proteomics*. **7**, 2318-29.
- A. Sundquist, M. Ronaghi, H. Tang, P. Pevzner, S. Batzoglou (2007) Whole-genome sequencing and assembly with high-throughput, short-read technologies. *PLoS ONE* **2**, e484.
- N. Bandeira, K.R. Clauser, P.A. Pevzner (2007) Shotgun protein sequencing: assembly of peptide tandem mass spectra from mixtures of modified proteins. *Mol Cell Proteomics*. **6**, 1123-34.
- N. Bandeira, D. Tsur, A. Frank, P.A. Pevzner (2007) Protein identification by spectral networks analysis. *Proc Natl Acad Sci U S A*. **104**, 6140-5
- M.A. Alekseyev, P.A. Pevzner (2007) Colored de Bruijn graphs and the genome halving problem. *IEEE/ACM Trans Comput Biol Bioinform.* **4**, 98-107.
- D. Zhi, U. Keich, P. Pevzner, S. Heber, H. Tang (2007) Correcting base-assignment errors in repeat regions of shotgun assembly. *IEEE/ACM Trans Comput Biol Bioinform.* **4**, 54-64.
- M.A. Alekseyev, P.A. Pevzner (2007) Whole genome duplications, multi-break rearrangements, and genome halving problem. *ACM/SIAM Symposium on Discrete Algorithms (SODA 2007)*, New Orleans, 665-679
- M. A. Alekseyev, P. A. Pevzner (2007) Whole Genome Duplications and Contracted Breakpoint Graphs. *SIAM J. Comput.*, **36**, 1748-1763
- Z. Jiang, H. Tang, M. Ventura, M.F. Cardone, T. Marques-Bonet, X. She, P.A. Pevzner, E.E. Eichler

(2007) Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. *Nature Genet.*

Books

A. Alexandrov, N. Alexandrov, M. Borodovsky, Y. Kalambet, A. Kister, A. Mironov, P.A. Pevzner, V. Shepelev. *Computer Analysis of Genetic Texts*, Nauka, Moscow, 1990 (in Russian)

P.A. Pevzner. *Computational Molecular Biology: An Algorithmic Approach*, The MIT Press, Cambridge, MA 2000

N.C. Jones, P.A. Pevzner. *Introduction to Bioinformatics Algorithms*. The MIT Press, Cambridge, MA 2004

Degrees Supervised

- | | |
|------|---|
| 1994 | Ph.D. Vineet Bafna. Combinatorial Methods in Molecular Evolution.
(currently Assistant Professor at UCSD) |
| 1995 | Ph.D. Sridhar Hannenhalli. Genome Rearrangements
(currently Assistant Professor at University of Pennsylvania) |
| 2000 | Ph.D. Sing Hoi Sze. Pattern Discovery in DNA sequences.
(currently Assistant Professor at Texas A&M University) |
| 2000 | Ph.D. Zufar Mulyukov. Computational Mass-spectrometry
(currently at Ceres, Inc) |
| 2001 | Ph.D. Earl Hubbell. Combinatorial problems related to DNA arrays.
(currently Senior Bioinformatician at Affymetrix) |
| 2002 | Ph.D. Guillaume Bourque. Multiple Genome Rearrangements.
(currently Director of Bioinformatics at Genomics Institute of Singapore) |
| 2005 | Ph.D. Degui Zhi. Analysis of mosaic structure in biomolecular sequences.
(currently postdoc at University of California at Berkeley) |
| 2007 | Ph.D. Neal Jones. Comparative genomics approach to motif finding.
(currently at Google) |
| 2007 | Ph.D. Nuno Bandeira. Spectral Networks and Shotgun Protein Sequencing.
(currently postdoc at University of California at San Diego) |
| 2007 | Ph.D. Max Alekseyev. Genome Rearrangements and Duplications.
(currently postdoc at University of California at San Diego) |

Former Postdoctoral Researchers and Project Leaders

- | | |
|-----------|--|
| 2000-2003 | Uri Keich (currently Assistant Professor at Cornell University) |
| 2001-2003 | Glenn Tesler (currently Assistant Professor at UCSD) |
| 2001-2003 | Steffen Heber (currently Assistant Professor at North Carolina State University) |
| 2001-2004 | Haixu Tang (currently Assistant Professor at University of Indiana) |
| 2003-2005 | Alkes Price (currently Research Scientist at Harvard Medical School) |