

Educational papers in Bioinformatics

Links to the journals have been provided. Access to papers is restricted to subscribers. Compiled by Xianjun Dong (xdong@rics.bwh.harvard.edu).

Nature Biotechnology Primers

<http://www.nature.com/nbt/collections/compbio/index.html>

Last Update: 2017-07-20

What is flux balance analysis?

Jeffrey D Orth, Ines Thiele and Bernhard θ Palsson

doi:10.1038/nbt.1614

Nature Biotechnology 28, 245-248 (2010)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is a support vector machine?

William S Noble

doi:10.1038/nbt1206-1565

Nature Biotechnology 24, 1565-1567 (2006)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is principal component analysis?

Markus Ringnér

doi:10.1038/nbt0308-303

Nature Biotechnology 26, 303-304 (2008)

[Abstract](#) | [Full Text](#) | [PDF](#)

How does gene expression clustering work?

Patrik D'haeseleer

doi:10.1038/nbt1205-1499

Nature Biotechnology 23, 1499-1501 (2005)

[Abstract](#) | [Full Text](#) | [PDF](#)

How to map billions of short reads onto genomes

Cole Trapnell and Steven L Salzberg

doi:10.1038/nbt0509-455

Nature Biotechnology 27, 455-457 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is a hidden Markov model?

Sean R Eddy

doi:10.1038/nbt1004-1315

Nature Biotechnology 22, 1315-1316 (2004)

[Abstract](#) | [Full Text](#) | [PDF](#)

How does multiple testing correction work?

William S Noble

doi:10.1038/nbt1209-1135

Nature Biotechnology 27, 1135-1137 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

Where did the BLOSUM62 alignment score matrix come from?

Sean R Eddy

doi:10.1038/nbt0804-1035

Nature Biotechnology 22, 1035-1036 (2004)

[Abstract](#) | [Full Text](#) | [PDF](#)

What are DNA sequence motifs?

Patrik D'haeseleer

doi:10.1038/nbt0406-423

Nature Biotechnology 24, 423-425 (2006)

[Abstract](#) | [Full Text](#) | [PDF](#)

How to apply de Bruijn graphs to genome assembly

Phillip E C Compeau, Pavel A Pevzner and Glenn Tesler

doi:10.1038/nbt.2023

Nature Biotechnology 29, 987-991 (2011)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is the expectation maximization algorithm?

Chuong B Do and Serafim Batzoglou

doi:10.1038/nbt1406

Nature Biotechnology 26, 897-899 (2008)

[Abstract](#) | [Full Text](#) | [PDF](#)

What are artificial neural networks?

Anders Krogh

doi:10.1038/nbt1386

Nature Biotechnology 26, 195-197 (2008)

[Abstract](#) | [Full Text](#) | [PDF](#)

How do RNA folding algorithms work?

Sean R Eddy

doi:10.1038/nbt1104-1457

Nature Biotechnology 22, 1457-1458 (2004)

[Abstract](#) | [Full Text](#) | [PDF](#)

How do shotgun proteomics algorithms identify proteins?

Edward M Marcotte

doi:10.1038/nbt0707-755

Nature Biotechnology 25, 755-757 (2007)

[Abstract](#) | [Full Text](#) | [PDF](#)

Inference in Bayesian networks

Chris J Needham, James R Bradford, Andrew J Bulpitt and David R Westhead

doi:10.1038/nbt0106-51

Nature Biotechnology 24, 51-53 (2006)

[Abstract](#) | [Full Text](#) | [PDF](#)

What are decision trees?

Carl Kingsford and Steven L Salzberg

doi:10.1038/nbt0908-1011

Nature Biotechnology 26, 1011-1013 (2008)

[Abstract](#) | [Full Text](#) | [PDF](#)

How does DNA sequence motif discovery work?

Patrik D'haeseleer

doi:10.1038/nbt0806-959

Nature Biotechnology 24, 959-961 (2006)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is dynamic programming?

Sean R Eddy

doi:10.1038/nbt0704-909

Nature Biotechnology 22, 909-910 (2004)

[Abstract](#) | [Full Text](#) | [PDF](#)

What is Bayesian statistics?

Sean R Eddy

doi:10.1038/nbt0904-1177

Nature Biotechnology 22, 1177-1178 (2004)

[Abstract](#) | [Full Text](#) | [PDF](#)

How to visually interpret biological data using networks

Daniele Merico, David Gfeller and Gary D Bader

doi:10.1038/nbt.1567

Nature Biotechnology 27, 921-924 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

SNP imputation in association studies

Eran Halperin and Dietrich A Stephan

doi:10.1038/nbt0409-349

Nature Biotechnology 27, 349-351 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

Analyzing 'omics data using hierarchical models

Hongkai Ji and X Shirley Liu

doi:10.1038/nbt.1619

Nature Biotechnology 28, 337-340 (2010)

[Abstract](#) | [Full Text](#) | [PDF](#)

How does eukaryotic gene prediction work?

Michael R Brent

doi:10.1038/nbt0807-883

Nature Biotechnology 25, 883-885 (2007)

[Abstract](#) | [Full Text](#) | [PDF](#)

Understanding genome browsing

Melissa S Cline and W James Kent

doi:10.1038/nbt0209-153

Nature Biotechnology 27, 153-155 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

Maximizing power in association studies

Eran Halperin and Dietrich A Stephan

doi:10.1038/nbt0309-255

Nature Biotechnology 27, 255-256 (2009)

[Abstract](#) | [Full Text](#) | [PDF](#)

Nature Methods: Points of Significance

<https://www.nature.com/search?journal=nmeth&title=%22points%20of%20significance%3A%22>

Last update: 2017-07-20

29 June 2017

[Principal component analysis](#)

Nature Methods 14, 641–642

30 May 2017

[Clustering](#)

Naomi Altman & Martin Krzywinski, Nature Methods 14, 545–546

30 March 2017

[Tabular data](#)

Naomi Altman & Martin Krzywinski, Nature Methods 14, 329–330

28 February 2017

[Interpreting P values](#)

Naomi Altman & Martin Krzywinski, Nature Methods 14, 213–214

29 December 2016

[P values and the search for significance](#)

Naomi Altman & Martin Krzywinski, Nature Methods 14, 3–4

29 September 2016

[Regularization](#)

Jake Lever, Martin Krzywinski & Naomi Altman, Nature Methods 13, 803–804

30 August 2016

[Model selection and overfitting](#)

Jake Lever, Martin Krzywinski & Naomi Altman, Nature Methods 13, 703–704

28 July 2016

[Classification evaluation](#)

Jake Lever, Martin Krzywinski & Naomi Altman, Nature Methods 13, 603–604

29 June 2016

[Logistic regression](#)

Jake Lever, Martin Krzywinski & Naomi Altman, Nature Methods 13, 541–542

28 April 2016

[Regression diagnostics](#)

Naomi Altman & Martin Krzywinski, Nature Methods 13, 385–386

30 March 2016

[Analyzing outliers: influential or nuisance?](#)

Naomi Altman & Martin Krzywinski, Nature Methods 13, 281–282

01 December 2015

[Multiple linear regression](#)

Martin Krzywinski & Naomi Altman, Nature Methods 12, 1103–1104

29 October 2015

[Simple linear regression](#)

Naomi Altman & Martin Krzywinski, Nature Methods 12, 999–1000

29 September 2015

[Association, correlation and causation](#)

Naomi Altman & Martin Krzywinski, Nature Methods 12, 899–900

28 August 2015

[Bayesian networks](#)

Jorge López Puga, Martin Krzywinski & Naomi Altman, Nature Methods 12, 799–800

28 May 2015

[Sampling distributions and the bootstrap](#)

Anthony Kulesa, Martin Krzywinski, Naomi Altman, Nature Methods 12, 477–478

29 April 2015

[Bayesian statistics](#)

Jorge López Puga, Martin Krzywinski & Naomi Altman, Nature Methods 12, 377–378

31 March 2015

[Bayes' theorem](#)

Jorge López Puga, Martin Krzywinski & Naomi Altman, Nature Methods 12, 277–278

26 February 2015

[Split plot design](#)

Naomi Altman & Martin Krzywinski, Nature Methods 12, 165–166

30 December 2014

[Sources of variation](#)

Naomi Altman & Martin Krzywinski, Nature Methods 12, 5–6

25 November 2014

[Two-factor designs](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 1187–1188

29 September 2014

[Nested designs](#)

Martin Krzywinski, Naomi Altman & Paul Blainey, Nature Methods 11, 977–978

28 August 2014

[Replication](#)

Paul Blainey, Martin Krzywinski & Naomi Altman, Nature Methods 11, 879–880

27 June 2014

[Analysis of variance and blocking](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 699–700

29 May 2014

[Designing comparative experiments](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 597–598

29 April 2014

[Nonparametric tests](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 467–468

28 March 2014

[Comparing samples---part II](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 355–356

27 February 2014

[Comparing samples—part I](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 215–216

30 January 2014

[Visualizing samples with box plots](#)

Martin Krzywinski & Naomi Altman, Nature Methods 11, 119–120

26 November 2013

[Power and sample size](#)

Martin Krzywinski & Naomi Altman, Nature Methods 10, 1139–1140

30 October 2013

[Significance, P values and t-tests](#)

Martin Krzywinski & Naomi Altman, Nature Methods 10, 1041–1042

27 September 2013

[Error bars](#)

Martin Krzywinski & Naomi Altman, Nature Methods 10, 921–922

29 August 2013

[Importance of being uncertain](#)

Martin Krzywinski & Naomi Altman, Nature Methods 10, 809–810

Nature Methods: Points of View

<https://www.nature.com/search?journal=nmeth&title=%22points%20of%20view%3A%22>

Last update: 2017-07-20

01 August 2010

[Color coding](#)

Bang Wong, Nature Methods 7, 573

01 September 2010

[Design of data figures](#)

Bang Wong, Nature Methods 7, 665

29 September 2010

[Salience](#)

Bang Wong, Nature Methods 7, 773

28 October 2010

[Gestalt principles \(Part 1\)](#)

Bang Wong, Nature Methods 7, 863

29 November 2010

[Gestalt principles \(Part 2\)](#)

Bang Wong, Nature Methods 7, 941

28 January 2011

[Points of review \(part 1\)](#)

Bang Wong, Nature Methods 8, 101

25 February 2011

[Points of review \(part 2\)](#)

Bang Wong, Nature Methods 8, 189

30 March 2011

[Typography](#)

Bang Wong, Nature Methods 8, 277

28 April 2011

[The overview figure](#)

Bang Wong, Nature Methods 8, 365

27 May 2011

[Color blindness](#)

Bang Wong, Nature Methods 8, 441

29 June 2011

[Avoiding color](#)

Bang Wong, Nature Methods 8, 525

28 July 2011

[Simplify to clarify](#)

Bang Wong, Nature Methods 8, 611

30 August 2011

[Arrows](#)

Bang Wong, Nature Methods 8, 701

29 September 2011

[Layout](#)

Bang Wong, Nature Methods 8, 783

28 October 2011

[Salience to relevance](#)

Bang Wong, Nature Methods 8, 889

29 November 2011

[The design process](#)

Bang Wong, Nature Methods 8, 987

28 December 2011

[Data exploration](#)

Noam Shoresh & Bang Wong, Nature Methods 9, 5

30 January 2012

[Networks](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 115

28 February 2012

[Heat maps](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 213

27 March 2012

[Integrating data](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 315

27 April 2012

[Representing the genome](#)

Cydney Nielsen & Bang Wong, Nature Methods 9, 423

30 May 2012

[Managing deep data in genome browsers](#)

Cydney Nielsen & Bang Wong, Nature Methods 9, 521

28 June 2012

[Representing genomic structural variation](#)

Cydney Nielsen & Bang Wong, Nature Methods 9, 631

30 July 2012

[Mapping quantitative data to color](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 769

30 August 2012

[Into the third dimension](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 851

27 September 2012

[Power of the plane](#)

Nils Gehlenborg & Bang Wong, Nature Methods 9, 935

06 November 2012

[Pencil and paper](#)

Bang Wong & Rikke Schmidt Kjærgaard, Nature Methods 9, 1037

07 December 2012

[Visualizing biological data](#)

...for depicting scientific data. Although Points of View will go on hiatus after...

Bang Wong, Nature Methods 9, 1131

27 February 2013

[Axes, ticks and grids](#)

Martin Krzywinski, Nature Methods 10, 183

28 March 2013

[Labels and callouts](#)

Martin Krzywinski, Nature Methods 10, 275

29 April 2013

[Elements of visual style](#)

Martin Krzywinski, Nature Methods 10, 371

30 May 2013

[Plotting symbols](#)

Martin Krzywinski & Bang Wong, Nature Methods 10, 451

27 June 2013

[Multidimensional data](#)

Martin Krzywinski & Erica Savig, Nature Methods 10, 595

30 July 2013

[Storytelling](#)

Martin Krzywinski & Alberto Cairo, Nature Methods 10, 687

30 January 2014

[Bar charts and box plots](#)

Marc Streit & Nils Gehlenborg, Nature Methods 11, 117

30 July 2014

[Sets and intersections](#)

Alexander Lex & Nils Gehlenborg, Nature Methods 11, 779

29 January 2015

[Temporal data](#)

Marc Streit & Nils Gehlenborg, Nature Methods 12, 97

30 June 2015

[Unentangling complex plots](#)

Gregor McInerney & Martin Krzywinski, Nature Methods 12, 591

30 December 2015

[Pathways](#)

Barbara J Hunnicutt & Martin Krzywinski, Nature Methods 13, 5

25 February 2016

[Neural circuit diagrams](#)

Barbara Jeanine Hunnicutt & Martin Krzywinski, Nature Methods 13, 189

31 May 2016

[Binning high-resolution data](#)

Martin Krzywinski, Nature Methods 13, 463

31 October 2016

[Intuitive design](#)

Martin Krzywinski, Nature Methods 13, 895

Critical Care: Medical Statistics

<https://www.biomedcentral.com/collections/CC-Medical>