Chapter 2

Protocols and analysis of single-cell DNA methylation data

Our undestanding of DNA methylation has been revolutionized by the development of BS-seq, which offers single-cytosine resolution and absolute quantification of methyl-cytosine genome-wide. Recent advances have demonstrated the power of single-cell sequencing to deconvolve mixed cell populations [57, 28, 81]. Incorporating epigenetic information into this single-cell arsenal will provide insights into epigenetic heterogeneity and transform our understanding of gene regulation.

The first section of this chapter describes scBS-seq, a protocol for genome-wide profiling of DNA methylation in single cells, a statistical method to assess methylation heterogeneity between cells, and applications in embryonic stem cells. The work is based on Smallwood et al. [108], which was joint work of Sebastien Smallwood, Heather Lee, Christof Angermueller, Felix Krueger, Heba Saadeh, Julian Peat, Simon Andrews, Oliver Stegle, and Wolf Reik.

> Individual contributions: Sebastien Smallwood and Heather Lee designed the study, prepared scBS-seq libraries, analyzed data and wrote the manuscript. Felix Krueger, Heba Saadeh, and Sebastien Smallwood performed sequence mapping and analyzed data. Julian Peat contributed to technical developments. Christof Angermueller and Oliver Stegle analysed the data.

The second section describes the scM&T-seq protocol for parallel profiling of DNA methylation and gene expression in single cells, methods to quantify associations between DNA

methylation and gene expression, and applications to mouse embryonic stem cells. The work is based on Angermueller et al. [6], which was joint work of Christof Angermueller, Stephen Clark, Heather Lee, Iain Macaulay, Mabel Teng, Tim Xiaoming Hu, Felix Krueger, Sebastien Smallwood, Chris Ponting, Thierry Voet, Gavin Kelsey, Oliver Stegle, and Wolf Reik.

Individual contributions: Christof Angermueller performed all statistical analyses of the data. Heather Lee, Iain Macaulay, Stephen Clark, and Sebastien Smallwood developed the protocol and performed experiments. Heather Lee, Iain Macaulay, Christof Angermueller, Stephen Clark, Oliver Stegle, Wolf Reik, and Chris Ponting interpreted the results. Mabel Teng contributed to method development. Tim Xiaoming Hu processed RNA-seq data. Felix Krueger processed BS-seq data. Wolf Reik, Gavin Kelsey, Iain Macaula, and Thierry Voet contributed protocols and reagents. Heather Lee, Iain Macaulay, Wolf Reik, and Thierry Voet conceived the project.

References

[1] Martın Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, and Matthieu Devin. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.

[2] Michalis Agathocleous, Georgia Christodoulou, Vasilis Promponas, Chris Christodoulou, Vassilis Vassiliades, and Antonis Antoniou. Protein secondary structure prediction with bidirectional recurrent neural nets: Can weight updating for each residue enhance performance? In *Artificial Intelligence Applications and Innovations*, pages 128–137. Springer. ISBN 3-642-16238-X.

[3] Guillaume Alain, Yoshua Bengio, and Salah Rifai. Regularized auto-encoders estimate local statistics. pages 1–17.

[4] F. W. Albert, S. Treusch, A. H. Shockley, J. S. Bloom, and L. Kruglyak. Genetics of single-cell protein abundance variation in large yeast populations. 506(7489):494–7. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature12904.

[5] Babak Alipanahi, Andrew Delong, Matthew T. Weirauch, and Brendan J. Frey. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. 33:831–838. ISSN 1087-0156. doi: 10.1038/nbt.3300.

[6] Christof Angermueller, Stephen J. Clark, Heather J. Lee, Iain C. Macaulay, Mabel J. Teng, Tim Xiaoming Hu, Felix Krueger, Sébastien A. Smallwood, Chris P. Ponting, Thierry Voet, Gavin Kelsey, Oliver Stegle, and Wolf Reik. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. 13(3):229–232, . ISSN 1548-7091. doi: 10.1038/nmeth.3728. URL http://www.nature.com/nmeth/journal/v13/n3/full/nmeth.3728.html.

[7] Christof Angermueller, Heather Lee, Wolf Reik, and Oliver Stegle. Accurate prediction of single-cell DNA methylation states using deep learning. page 055715, . doi: 10.1101/055715. URL http://biorxiv.org/content/early/2017/02/01/055715.

[8] Ehsaneddin Asgari and Mohammad R. K. Mofrad. ProtVec: A Continuous Distributed Representation of Biological Sequences. 10:e0141287. ISSN 1932-6203. doi: 10.1371/journal.pone.0141287.

[9] S. Bach, A. Binder, G. Montavon, F. Klauschen, K. R. Muller, and W. Samek. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance

Propagation. 10(7):e0130140. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0130140.

[10] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate.

[11] Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, James Bergstra, Ian Goodfellow, Arnaud Bergeron, Nicolas Bouchard, David Warde-Farley, and Yoshua Bengio. Theano: New features and speed improvements.

[12] A. Battle, Z. Khan, S. H. Wang, A. Mitrano, M. J. Ford, J. K. Pritchard, and Y. Gilad. Genomic variation. Impact of regulatory variation from RNA to protein. 347(6222): 664–7. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1260793.

[13] Jordana T Bell, Athma A Pai, Joseph K Pickrell, Daniel J Gaffney, Roger Pique-Regi, Jacob F Degner, Yoav Gilad, and Jonathan K Pritchard. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. 12(1):R10.

[14] Yoshua Bengio. Practical recommendations for gradient-based training of deep architectures. In *Neural Networks: Tricks of the Trade*, pages 437–478. Springer. ISBN 3-642-35288-X.

[15] Yoshua Bengio, Aaron Courville, and Pierre Vincent. Representation learning: A review and new perspectives. 35(8):1798–1828, . ISSN 0162-8828.

[16] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy layer-wise training of deep networks. volume 19, page 153, . ISBN 1049-5258.

[17] James Bergstra and David D. Cox. Hyperparameter Optimization and Boosting for Classifying Facial Expressions: How good can a "Null" Model be?

[18] Zhengping Che, Sanjay Purushotham, Robinder Khemani, and Yan Liu. Distilling Knowledge from Deep Networks with Applications to Healthcare Domain.

[19] C. Cheng, K. K. Yan, K. Y. Yip, J. Rozowsky, R. Alexander, C. Shou, and M. Gerstein. A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets. 12(2):R15. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/gb-2011-12-2-r15.

[20] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling.

[21] Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. pages 2843–2851.

[22] Dan C Cireşan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Mitosis detection in breast cancer histology images with deep neural networks. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 411–418. Springer. ISBN 3-642-40762-5.

[23] Ronan Collobert, Koray Kavukcuoglu, and Clément Farabet. Torch7: A matlab-like environment for machine learning. EPFL-CONF-192376.

[24] George E. Dahl, Navdeep Jaitly, and Ruslan Salakhutdinov. Multi-task Neural Networks for QSAR Predictions.

[25] Yann N Dauphin, Razvan Pascanu, Caglar Gulcehre, Kyunghyun Cho, Surya Ganguli, and Yoshua Bengio. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. pages 2933–2941.

[26] Li Deng. Deep Learning: Methods and Applications. 7:197–387. ISSN 1932-8346, 1932-8354. doi: 10.1561/2000000039.

[27] Li Deng and Roberto Togneri. Deep dynamic models for learning hidden representations of speech features. In *Speech and Audio Processing for Coding, Enhancement and Recognition*, pages 153–195. Springer.

[28] Q. Deng, D. Ramskold, B. Reinius, and R. Sandberg. Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. 343(6167): 193–196. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1245316. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1245316.

[29] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition.

[30] John Duchi, Elad Hazan, and Yoram Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. 12:2121–2159. ISSN ISSN 1533-7928. URL http://www.jmlr.org/papers/v12/duchi11a.html.

[31] F. Eduati, L. M. Mangravite, T. Wang, H. Tang, J. C. Bare, R. Huang, T. Norman, M. Kellen, M. P. Menden, J. Yang, X. Zhan, R. Zhong, G. Xiao, M. Xia, N. Abdo, O. Kosyk, Niehs-Ncats-Unc Dream Toxicogenetics Collaboration, S. Friend, A. Dearry, A. Simeonov, R. R. Tice, I. Rusyn, F. A. Wright, G. Stolovitzky, Y. Xie, and J. Saez-Rodriguez. Prediction of human population responses to toxic compounds by a collaborative competition. 33(9):933–40. ISSN 1546-1696 (Electronic) 1087-0156 (Linking). doi: 10.1038/nbt.3299.

[32] Jesse Eickholt and Jianlin Cheng. DNdisorder: Predicting protein disorder using boosting and deep networks. 14:88, . ISSN 1471-2105. doi: 10.1186/1471-2105-14-88.

[33] Jesse Eickholt and Jianlin Cheng. Predicting protein residue-residue contacts using deep networks and boosting. 28:3066–3072, . ISSN 1367-4803. doi: 10.1093/bioinformatics/bts598.

[34] Rasool Fakoor, Faisal Ladhak, Azade Nazi, and Manfred Huber. Using deep learning to enhance cancer diagnosis and classification.

[35] BWAC Farley and W Clark. Simulation of self-organizing systems by digital computer. 4(4):76–84. ISSN 2168-2690.

[36] Alessandro Ferrari, Stefano Lombardi, and Alberto Signoroni. Bacterial colony counting by Convolutional Neural Networks. pages 7458–7461. IEEE.

[37] Erik Gawehn, Jan A. Hiss, and Gisbert Schneider. Deep Learning in Drug Discovery. 35(1):3–14. ISSN 1868-1751. doi: 10.1002/minf.201501008.

[38] J Raphael Gibbs, Marcel P van der Brug, Dena G Hernandez, Bryan J Traynor, Michael A Nalls, Shiao-Lin Lai, Sampath Arepalli, Allissa Dillman, Ian P Rafferty, and Juan Troncoso. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. 6(5):e1000952. ISSN 1553-7404.

[39] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. pages 580–587.

[40] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. pages 249–256.

[41] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. pages 315–323.

[42] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. URL http://www.deeplearningbook.org.

[43] A. Graves, A.-R. Mohamed, and G. Hinton. Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6645–6649. doi: 10.1109/ICASSP.2013.6638947.

[44] Alex Graves. Generating Sequences With Recurrent Neural Networks.

[45] F. Grubert, J. B. Zaugg, M. Kasowski, O. Ursu, D. V. Spacek, A. R. Martin, P. Greenside, R. Srivas, D. H. Phanstiel, A. Pekowska, N. Heidari, G. Euskirchen, W. Huber, J. K. Pritchard, C. D. Bustamante, L. M. Steinmetz, A. Kundaje, and M. Snyder. Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. 162(5):1051–65. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2015.07.048.

[46] Trevor Hastie, Robert Tibshirani, Jerome Friedman, and James Franklin. The elements of statistical learning: Data mining, inference and prediction. 27(2):83–85. ISSN 0343-6993.

[47] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. .

[48] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. pages 1026–1034, .

[49] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, and Tara N Sainath. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. 29(6):82–97. ISSN 1053-5888.

[50] Geoffrey E Hinton. A practical guide to training restricted boltzmann machines. In *Neural Networks: Tricks of the Trade*, pages 599–619. Springer. ISBN 3-642-35288-X.

[51] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. 313(5786):504–507. ISSN 0036-8075.

[52] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. 9(8):1735–1780. ISSN 0899-7667. doi: 10.1162/neco.1997.9.8.1735. URL http://dx.doi.org/10.1162/neco.1997.9.8.1735.

[53] David H Hubel and Torsten N Wiesel. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. 206(2):419, .

[54] DH Hubel and TN Wiesel. Shape and arrangement of columns in cat's striate cortex. 165(3):559, .

[55] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *Learning and Intelligent Optimization*, pages 507–523. Springer. ISBN 3-642-25565-5.

[56] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.

[57] Viren Jain, Joseph F Murray, Fabian Roth, Srinivas Turaga, Valentin Zhigulin, Kevin L Briggman, Moritz N Helmstaedter, Winfried Denk, and H Sebastian Seung. Supervised learning of image restoration with convolutional networks. pages 1–8. IEEE. ISBN 1-4244-1630-2.

[58] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun. What is the best multi-stage architecture for object recognition? In *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153. doi: 10.1109/ICCV.2009.5459469.

[59] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. pages 675–678. ACM. ISBN 1-4503-3063-0.

[60] H. M. Kang, C. Ye, and E. Eskin. Accurate discovery of expression quantitative trait loci under confounding from spurious and genuine regulatory hotspots. 180(4):1909–25. ISSN 0016-6731 (Print) 0016-6731 (Linking). doi: 10.1534/genetics.108.094201.

[61] R. Karlic, H. R. Chung, J. Lasserre, K. Vlahovicek, and M. Vingron. Histone modification levels are predictive for gene expression. 107(7):2926–31. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.0909344107.

[62] Andrej Karpathy and Justin Johnson. CS231n Convolutional Neural Networks for Visual Recognition. URL http://cs231n.github.io/.

[63] D. B. Kell. Metabolomics, machine learning and modelling: Towards an understanding of the language of cells. 33:520–4. ISSN 0300-5127 (Print) 0300-5127 (Linking). doi: 10.1042/BST0330520.

[64] D. R. Kelley, J. Snoek, and J. Rinn. Basset: Learning the regulatory code of the accessible genome with deep convolutional neural networks. Advance online. ISSN 1549-5469 (Electronic) 1088-9051 (Linking). doi: 10.1101/gr.200535.115.

[65] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization.

[66] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes.

[67] Pang Wei Koh, Emma Pierson, and Anshul Kundaje. Denoising genome-wide histone ChIP-seq with convolutional neural networks. page 052118. doi: 10.1101/052118. URL http://biorxiv.org/content/early/2017/01/27/052118.

[68] Oren Z Kraus, Lei Jimmy Ba, and Brendan Frey. Classifying and Segmenting Microscopy Images Using Convolutional Multiple Instance Learning.

[69] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. pages 1097–1105.

[70] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. 521:436–444, . ISSN 0028-0836. doi: 10.1038/nature14539.

[71] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. 1(4):541–551, .

[72] Byunghan Lee, Taehoon Lee, Byunggook Na, and Sungroh Yoon. DNA-Level Splice Junction Prediction using Deep Recurrent Neural Networks.

[73] Michael K. K. Leung, Hui Yuan Xiong, Leo J. Lee, and Brendan J. Frey. Deep learning of the tissue-regulated splicing code. 30:i121–i129, . ISSN 1367-4803, 1460-2059. doi: 10.1093/bioinformatics/btu277.

[74] Michael KK Leung, Andrew Delong, Babak Alipanahi, and Brendan J Frey. Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. . ISSN 0018-9219.

[75] J. Li, T. Ching, S. Huang, and L. X. Garmire. Using epigenomics data to predict gene expression in lung cancer. 16 Suppl 5:S10. ISSN 1471-2105 (Electronic) 1471-2105 (Linking). doi: 10.1186/1471-2105-16-S5-S10.

[76] Stan Z Li. *Markov Random Field Modeling in Image Analysis*. Springer Science & Business Media. ISBN 1-84800-279-3.

[77] Zhen Li and Yizhou Yu. Protein Secondary Structure Prediction Using Cascaded Convolutional and Recurrent Neural Networks. URL http://arxiv.org/abs/1604.07176.

[78] M. W. Libbrecht and W. S. Noble. Machine learning applications in genetics and genomics. 16(6):321–32. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg3920.

[79] Zachary C. Lipton. A Critical Review of Recurrent Neural Networks for Sequence Learning.

[80] James Lyons, Abdollah Dehzangi, Rhys Heffernan, Alok Sharma, Kuldip Paliwal, Abdul Sattar, Yaoqi Zhou, and Yuedong Yang. Predicting backbone C-alpha angles and dihedrals from protein sequences by stacked sparse auto-encoder deep neural network. 35(28):2040–2046. ISSN 1096-987X. doi: 10.1002/jcc.23718. URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.23718/abstract.

[81] Iain C. Macaulay and Thierry Voet. Single Cell Genomics: Advances and Future Perspectives. 10(1):e1004126. ISSN 1553-7404. doi: 10.1371/journal.pgen.1004126. URL http://dx.plos.org/10.1371/journal.pgen.1004126.

[82] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov. Applications of Deep Learning in Biomedicine. 13(5):1445–54. ISSN 1543-8392 (Electronic) 1543-8384 (Linking). doi: 10.1021/acs.molpharmaceut.5b00982.

[83] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. 5(4):115–133. ISSN 0007-4985.

[84] M. P. Menden, F. Iorio, M. Garnett, U. McDermott, C. H. Benes, P. J. Ballester, and J. Saez-Rodriguez. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. 8(4):e61318. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0061318.

[85] Ryszard S Michalski, Jaime G Carbonell, and Tom M Mitchell. *Machine Learning: An Artificial Intelligence Approach*. Springer Science & Business Media. ISBN 3-662-12405-X.

[86] S. B. Montgomery, M. Sammeth, M. Gutierrez-Arcelus, R. P. Lach, C. Ingle, J. Nisbett, R. Guigo, and E. T. Dermitzakis. Transcriptome genetics using second generation sequencing in a Caucasian population. 464(7289):773–7. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature08903.

[87] Kevin P Murphy. *Machine Learning: A Probabilistic Perspective*. MIT press. ISBN 0-262-01802-0.

[88] K. Märtens, J. Hallin, J. Warringer, G. Liti, and L. Parts. Predicting quantitative traits from genome and phenome with near perfect accuracy. 7:11512. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/ncomms11512.

[89] Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*, volume 87. Springer Science & Business Media, . ISBN 1-4419-8853-X.

[90] Yurii Nesterov. A method of solving a convex programming problem with convergence rate O (1/k2). volume 27, pages 372–376, . 2.

[91] Feng Ning, Damien Delhomme, Yann LeCun, Fabio Piano, Léon Bottou, and Paolo Emilio Barbano. Toward automatic phenotyping of developing embryos from videos. 14(9):1360–1371. ISSN 1057-7149.

[92] Yongjin Park and Manolis Kellis. Deep learning for regulatory genomics. 33:825–826. ISSN 1087-0156. doi: 10.1038/nbt.3313.

[93] L. Parts, Y. C. Liu, M. M. Tekkedil, L. M. Steinmetz, A. A. Caudy, A. G. Fraser, C. Boone, B. J. Andrews, and A. P. Rosebrock. Heritability and genetic basis of protein level variation in an outbred population. 24(8):1363–70, . ISSN 1549-5469 (Electronic) 1088-9051 (Linking). doi: 10.1101/gr.170506.113.

[94] L. Parts, O. Stegle, J. Winn, and R. Durbin. Joint genetic analysis of gene expression data with inferred cellular phenotypes. 7(1):e1001276, . ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.1001276.

[95] J. K. Pickrell, J. C. Marioni, A. A. Pai, J. F. Degner, B. E. Engelhardt, E. Nkadori, J. B. Veyrieras, M. Stephens, Y. Gilad, and J. K. Pritchard. Understanding mechanisms underlying human gene expression variation with RNA sequencing. 464(7289):768–72. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature08872.

[96] T. Pärnamaa and L. Parts. Accurate classification of protein subcellular localization from high throughput microscopy images using deep learning. page 10.1101/050757. doi: 10.1101/050757.

[97] Daniel Quang and Xiaohui Xie. DanQ: A hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. 44(11): e107–e107. ISSN 0305-1048, 1362-4962. doi: 10.1093/nar/gkw226. URL http://nar.oxfordjournals.org/content/44/11/e107.

[98] B. Rakitsch and O. Stegle. Modelling local gene networks increases power to detect trans-acting genetic effects on gene expression. 17(1):33. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/s13059-016-0895-2.

[99] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pages 234–241. Springer. ISBN 3-319-24573-2.

[100] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. 65(6):386. ISSN 1939-1471.

[101] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. 5(3):1.

[102] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, and Michael Bernstein. Imagenet large scale visual recognition challenge. 115(3):211–252. ISSN 0920-5691.

[103] Ruslan Salakhutdinov and Geoffrey Hinton. An efficient learning procedure for deep Boltzmann machines. 24(8):1967–2006. ISSN 0899-7667.

[104] Ruslan Salakhutdinov and Hugo Larochelle. Efficient learning of deep Boltzmann machines. pages 693–700.

[105] Juergen Schmidhuber. Deep Learning in Neural Networks: An Overview. 61:85–117. ISSN 08936080. doi: 10.1016/j.neunet.2014.09.003.

[106] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition.

[107] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.

[108] Sébastien A. Smallwood, Heather J. Lee, Christof Angermueller, Felix Krueger, Heba Saadeh, Julian Peat, Simon R. Andrews, Oliver Stegle, Wolf Reik, and Gavin Kelsey. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. 11(8):817–820. ISSN 1548-7091. doi: 10.1038/nmeth.3035. URL http://www.nature.com/nmeth/journal/v11/n8/full/nmeth.3035.html.

[109] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. pages 2951–2959.

[110] Matt Spencer, Jesse Eickholt, and Jianlin Cheng. A deep learning network approach to ab initio protein secondary structure prediction. 12(1):103–112. ISSN 1545-5963.

[111] Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller. Striving for Simplicity: The All Convolutional Net.

[112] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. 15:1929–1958.

[113] O. Stegle, L. Parts, R. Durbin, and J. Winn. A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. 6(5):e1000770. ISSN 1553-7358 (Electronic) 1553-734X (Linking). doi: 10.1371/journal.pcbi.1000770.

[114] Gary D Stormo, Thomas D Schneider, Larry Gold, and Andrzej Ehrenfeucht. Use of the 'Perceptron'algorithm to distinguish translational initiation sites in E. coli. 10(9): 2997–3011. ISSN 0305-1048.

[115] Ilya Sutskever. Training recurrent neural networks.

[116] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. pages 3104–3112.

[117] A. L. Swan, A. Mobasheri, D. Allaway, S. Liddell, and J. Bacardit. Application of machine learning to proteomics data: Classification and biomarker identification in postgenomics biology. 17(12):595–610. ISSN 1557-8100 (Electronic) 1536-2310 (Linking). doi: 10.1089/omi.2013.0017.

[118] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision.

[119] Søren Kaae Sønderby and Ole Winther. Protein Secondary Structure Prediction with Long Short Term Memory Networks.

[120] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. 11:3371–3408. ISSN 1532-4435.

[121] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. pages 3156–3164.

[122] Kun Wang, Kan Cao, and Sridhar Hannenhalli. Chromatin and genomic determinants of alternative splicing. pages 345–354. ACM. ISBN 1-4503-3853-4.

[123] S. M. Waszak, O. Delaneau, A. R. Gschwind, H. Kilpinen, S. K. Raghav, R. M. Witwicki, A. Orioli, M. Wiederkehr, N. I. Panousis, A. Yurovsky, L. Romano-Palumbo, A. Planchon, D. Bielser, I. Padioleau, G. Udin, S. Thurnheer, D. Hacker, N. Hernandez, A. Reymond, B. Deplancke, and E. T. Dermitzakis. Population Variation and Genetic Control of Modular Chromatin Architecture in Humans. 162(5):1039–50. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2015.08.001.

[124] Yuanpu Xie, Fuyong Xing, Xiangfei Kong, Hai Su, and Lin Yang. Beyond Classification: Structured Regression for Robust Cell Detection Using Convolutional Neural Network. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pages 358–365. Springer. ISBN 3-319-24573-2.

[125] Caiming Xiong, Stephen Merity, and Richard Socher. Dynamic Memory Networks for Visual and Textual Question Answering. .

[126] Hui Y. Xiong, Babak Alipanahi, Leo J. Lee, Hannes Bretschneider, Daniele Merico, Ryan K. C. Yuen, Yimin Hua, Serge Gueroussov, Hamed S. Najafabadi, Timothy R. Hughes, Quaid Morris, Yoseph Barash, Adrian R. Krainer, Nebojsa Jojic, Stephen W. Scherer, Benjamin J. Blencowe, and Brendan J. Frey. The human splicing code reveals new insights into the genetic determinants of disease. 347:1254806, . ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1254806.

[127] Huijuan Xu and Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. URL http://arxiv.org/abs/1511.05234.

[128] Yan Xu, Tao Mo, Qiwei Feng, Peilin Zhong, Maode Lai, and Eric I Chang. Deep learning of feature representation with multiple instance learning for medical image analysis. pages 1626–1630. IEEE.

[129] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 3320–3328. Curran Associates, Inc.

[130] Matthew D. Zeiler and Rob Fergus. Visualizing and Understanding Convolutional Networks. In *Computer Vision – ECCV 2014*, pages 818–833. Springer, Cham. doi: 10.1007/978-3-319-10590-1_53. URL https://link.springer.com/chapter/10.1007/978-3-319-10590-1_53.

[131] Wenlu Zhang, Rongjian Li, Tao Zeng, Qian Sun, Sudhir Kumar, Jieping Ye, and Shuiwang Ji. Deep model based transfer and multi-task learning for biological image analysis. pages 1475–1484. ACM. ISBN 1-4503-3664-7.

[132] J. Zhou and O. G. Troyanskaya. Predicting effects of noncoding variants with deep learning-based sequence model. 12(10):931–4. ISSN 1548-7105 (Electronic) 1548-7091 (Linking). doi: 10.1038/nmeth.3547.