

1. Wei, X., Das, J., Fragoza, R., Liang, J., Bastos de Oliveira, F. M., Lee, H. R., Wang, X., Mort, M., Stenson, P. D., Cooper, D. N., Lipkin, S. M., Smolka, M. B., & Yu, H. (2014) A massively parallel pipeline to clone DNA variants and examine molecular phenotypes of human disease mutations.. *PLoS genetics* 10, e1004819.
2. Weinhold, N., Jacobsen, A., Schultz, N., Sander, C., & Lee, W. (2014) Genome-wide analysis of noncoding regulatory mutations in cancer.. *Nature genetics* 46, 1160--1165.
3. Fredriksson, N. J., Ny, L., Nilsson, J. A., & Larsson, E. (2014) Systematic analysis of noncoding somatic mutations and gene expression alterations across 14 tumor types.. *Nature genetics* 46, 1258--1263.
4. Fu, Y., Liu, Z., Lou, S., Bedford, J., Mu, X. J., Yip, K. Y., Khurana, E., & Gerstein, M. (2014) FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer.. *Genome biology* 15, 480.
5. Cruchaga, C., Karch, C. M., Jin, S. C., Benitez, B. A., Cai, Y., Guerreiro, R., Harari, O., Norton, J., Budde, J., Bertelsen, S., Jeng, A. T., Cooper, B., Skorupa, T., Carrell, D., Levitch, D., Hsu, S., Choi, J., Ryten, M., UK Brain Expression Consortium, Hardy, J., Trabzuni, D., Weale, M. E., Ramasamy, A., Smith, C., Sassi, C., Bras, J., Gibbs, J. R., Hernandez, D. G., Lupton, M. K., Powell, J., Forabosco, P., Ridge, P. G., Corcoran, C. D., Tschanz, J. T., Norton, M. C., Munger, R. G., Schmutz, C., Leary, M., Demirci, F. Y., Bamne, M. N., Wang, X., Lopez, O. L., Ganguli, M., Medway, C., Turton, J., Lord, J., Braae, A., Barber, I., Brown, K., Alzheimer's Research UK Consortium, Passmore, P., Craig, D., Johnston, J., McGuinness, B., Todd, S., Heun, R., Kölsch, H., Kehoe, P. G., Hooper, N. M., Vardy, E. R. L. C., Mann, D. M., Pickering-Brown, S., Kalsheker, N., Lowe, J., Morgan, K., David Smith, A., Wilcock, G., Warden, D., Holmes, C., Pastor, P., Lorenzo-Betancor, O., Brkanac, Z., Scott, E., Topol, E., Rogaeva, E., Singleton, A. B., Kamboh, M. I., St George-Hyslop, P., Cairns, N., Morris, J. C., Kauwe, J. S. K., & Goate, A. M. (2014) Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease.. *Nature* 505, 550--554.
6. MacArthur, D. G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J. K., Montgomery, S. B., Albers, C. A., Zhang, Z. D., Conrad, D. F., Lunter, G., Zheng, H., Ayub, Q., DePristo, M. A., Banks, E., Hu, M., Handsaker, R. E., Rosenfeld, J. A., Fromer, M., Jin, M., Mu, X. J., Khurana, E., Ye, K., Kay, M., Saunders, G. I., Suner, M.-M., Hunt, T., Barnes, I. H. A., Amid, C., Carvalho-Silva, D. R., Bignell, A. H., Snow, C., Yngvadottir, B., Bumpstead, S., Cooper, D. N., Xue, Y., Romero, I. G., 1000 Genomes Project Consortium, Wang, J., Li, Y., Gibbs, R. A., McCarroll, S. A., Dermitzakis, E. T., Pritchard, J. K., Barrett, J. C., Harrow, J., Hurles, M. E., Gerstein, M. B., & Tyler-Smith, C. (2012) A systematic survey of loss-of-function variants in human protein-coding genes.. *Science (New York, N.Y.)* 335, 823--828.
7. Cox, A., Dunning, A. M., Garcia-Closas, M., Balasubramanian, S., Reed, M. W. R., Pooley, K. A., Scollen, S., Baynes, C., Ponder, B. A. J., Chanock, S., Lissowska, J., Brinton, L., Peplonska, B., Southey, M. C., Hopper, J. L., McCredie, M. R. E., Giles, G. G., Fletcher, O., Johnson, N., dos Santos Silva, I., Gibson, L., Bojesen, S. E., Nordestgaard, B. G., Axelsson, C. K., Torres, D., Hamann, U., Justenhoven, C., Brauch, H., Chang-Claude, J., Kropp, S., Risch, A., Wang-Gohrke, S., Schürmann, P., Bogdanova, N., Dörk, T., Fagerholm, R., Aaltonen, K., Blomqvist, C., Nevanlinna, H., Seal, S., Renwick, A., Stratton, M. R., Rahman, N., Sangrajrang, S., Hughes, D., Odefrey, F., Brennan, P., Spurdle, A. B., Chenevix-Trench, G., Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer, Beesley, J., Mannermaa, A., Hartikainen, J., Kataja, V., Kosma, V.-M., Couch, F. J., Olson, J. E., Goode, E. L., Broeks, A., Schmidt, M. K., Hogervorst, F. B. L., Van't Veer, L. J., Kang, D., Yoo, K.-Y., Noh, D.-Y., Ahn, S.-H., Wedrén, S., Hall, P., Low, Y.-L., Liu, J., Milne, R. L., Ribas, G., Gonzalez-Neira, A., Benitez, J., Sigurdson, A. J., Stredrick, D. L., Alexander, B. H., Struewing, J. P., Pharoah, P. D. P., Easton, D. F., & Breast Cancer Association Consortium (2007) A common coding variant in CASP8 is associated with breast cancer risk..

Nature genetics 39, 352--358.

8. Momozawa, Y., Mni, M., Nakamura, K., Coppieters, W., Almer, S., Amininejad, L., Cleynen, I., Colombel, J.-F., de Rijk, P., Dewit, O., Finkel, Y., Gassull, M. A., Goossens, D., Laukens, D., Lémann, M., Libioulle, C., O'Morain, C., Reenaers, C., Rutgeerts, P., Tysk, C., Zelenika, D., Lathrop, M., Del-Favero, J., Hugot, J.-P., de Vos, M., Franchimont, D., Vermeire, S., Louis, E., & Georges, M. (2011) Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease.. *Nature genetics* 43, 43--47.
9. Kimchi-Sarfaty, C., Oh, J. M., Kim, I.-W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V., & Gottesman, M. M. (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity.. *Science (New York, N.Y.)* 315, 525--528.
10. Ward, L. D. & Kellis, M. (2012) Interpreting noncoding genetic variation in complex traits and human disease.. *Nature biotechnology* 30, 1095--1106.
11. De Gobbi, M., Viprakasit, V., Hughes, J. R., Fisher, C., Buckle, V. J., Ayyub, H., Gibbons, R. J., Vernimmen, D., Yoshinaga, Y., de Jong, P., Cheng, J.-F., Rubin, E. M., Wood, W. G., Bowden, D., & Higgs, D. R. (2006) A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter.. *Science (New York, N.Y.)* 312, 1215--1217.
12. Cheung, V. G. & Spielman, R. S. (2009) Genetics of human gene expression: mapping DNA variants that influence gene expression.. *Nature reviews. Genetics* 10, 595--604.
13. Li, Q., Seo, J.-H., Stranger, B., McKenna, A., Pe'er, I., Laframboise, T., Brown, M., Tyekucheva, S., & Freedman, M. L. (2013) Integrative eQTL-based analyses reveal the biology of breast cancer risk loci.. *Cell* 152, 633--641.
14. Huang, F. W., Hodis, E., Xu, M. J., Kryukov, G. V., Chin, L., & Garraway, L. A. (2013) Highly recurrent TERT promoter mutations in human melanoma.. *Science (New York, N.Y.)* 339, 957--959.
15. Horn, S., Figl, A., Rachakonda, P. S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., Schadendorf, D., & Kumar, R. (2013) TERT promoter mutations in familial and sporadic melanoma.. *Science (New York, N.Y.)* 339, 959--961.
16. Tournamille, C., Colin, Y., Cartron, J. P., & Le Van Kim, C. (1995) Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals.. *Nature genetics* 10, 224--228.
17. McCarroll, S. A., Huett, A., Kuballa, P., Chilewski, S. D., Landry, A., Goyette, P., Zody, M. C., Hall, J. L., Brant, S. R., Cho, J. H., Duerr, R. H., Silverberg, M. S., Taylor, K. D., Rioux, J. D., Altshuler, D., Daly, M. J., & Xavier, R. J. (2008) Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease.. *Nature genetics* 40, 1107--1112.
18. Verlaan, D. J., Ge, B., Grundberg, E., Hoberman, R., Lam, K. C. L., Koka, V., Dias, J., Gurd, S., Martin, N. W., Mallmin, H., Nilsson, O., Harmsen, E., Dewar, K., Kwan, T., & Pastinen, T. (2009) Targeted screening of cis-regulatory variation in human haplotypes.. *Genome research* 19, 118--127.
19. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., & Parkinson, H. (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations.. *Nucleic acids research* 42, D1001--D1006.
20. Ferrari, R., Hernandez, D. G., Nalls, M. A., Rohrer, J. D., Ramasamy, A., Kwok, J. B. J., Dobson-Stone, C., Brooks, W. S., Schofield, P. R., Halliday, G. M., Hodges, J. R., Pigué, O., Bartley, L., Thompson, E., Haan, E., Hernández, I., Ruiz, A., Boada, M., Borroni, B., Padovani, A., Cruchaga, C., Cairns, N. J., Benussi, L., Binetti, G., Ghidoni, R., Forloni, G., Galimberti, D., Fenoglio, C., Serpente, M., Scarpini, E., Clarimón, J., Lleó, A., Blesa, R., Waldö, M. L., Nilsson, K., Nilsson, C., Mackenzie, I. R. A., Hsiung, G.-Y. R., Mann, D. M. A., Grafman, J., Morris, C. M., Attems, J., Griffiths, T. D., McKeith, I. G., Thomas, A. J.,

Pietrini, P., Huey, E. D., Wassermann, E. M., Baborie, A., Jaros, E., Tierney, M. C., Pastor, P., Razquin, C., Ortega-Cubero, S., Alonso, E., Pernecky, R., Diehl-Schmid, J., Alexopoulos, P., Kurz, A., Rainero, I., Rubino, E., Pinessi, L., Rogaeva, E., St George-Hyslop, P., Rossi, G., Tagliavini, F., Giaccone, G., Rowe, J. B., Schlachetzki, J. C. M., Uphill, J., Collinge, J., Mead, S., Danek, A., Van Deerlin, V. M., Grossman, M., Trojanowski, J. Q., van der Zee, J., Deschamps, W., Van Langenhove, T., Cruts, M., Van Broeckhoven, C., Cappa, S. F., Le Ber, I., Hannequin, D., Golfier, V., Vercelletto, M., Brice, A., Nacmias, B., Sorbi, S., Bagnoli, S., Piaceri, I., Nielsen, J. E., Hjermand, L. E., Riemenschneider, M., Mayhaus, M., Ibach, B., Gasparoni, G., Pichler, S., Gu, W., Rossor, M. N., Fox, N. C., Warren, J. D., Spillantini, M. G., Morris, H. R., Rizzu, P., Heutink, P., Snowden, J. S., Rollinson, S., Richardson, A., Gerhard, A., Bruni, A. C., Maletta, R., Frangipane, F., Cupidi, C., Bernardi, L., Anfossi, M., Gallo, M., Conidi, M. E., Smirne, N., Rademakers, R., Baker, M., Dickson, D. W., Graff-Radford, N. R., Petersen, R. C., Knopman, D., Josephs, K. A., Boeve, B. F., Parisi, J. E., Seeley, W. W., Miller, B. L., Karydas, A. M., Rosen, H., van Swieten, J. C., Dopper, E. G. P., Seelaar, H., Pijnenburg, Y. A. L., Scheltens, P., Logroscino, G., Capozzo, R., Novelli, V., Puca, A. A., Franceschi, M., Postiglione, A., Milan, G., Sorrentino, P., Kristiansen, M., Chiang, H.-H., Graff, C., Pasquier, F., Rollin, A., Deramecourt, V., Lebert, F., Kapogiannis, D., Ferrucci, L., Pickering-Brown, S., Singleton, A. B., Hardy, J., & Momeni, P. (2014) Frontotemporal dementia and its subtypes: a genome-wide association study.. *The Lancet. Neurology* 13, 686--699.

21. Hindorf, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9362--9367.

22. Stacey, S. N., Manolescu, A., Sulem, P., Rafnar, T., Gudmundsson, J., Gudjonsson, S. A., Masson, G., Jakobsdottir, M., Thorlacius, S., Helgason, A., Aben, K. K., Strobbe, L. J., Albers-Akkers, M. T., Swinkels, D. W., Henderson, B. E., Kolonel, L. N., Le Marchand, L., Millastre, E., Andres, R., Godino, J., Garcia-Prats, M. D., Polo, E., Tres, A., Mouy, M., Saemundsdottir, J., Backman, V. M., Gudmundsson, L., Kristjansson, K., Bergthorsson, J. T., Kostic, J., Frigge, M. L., Geller, F., Gudbjartsson, D., Sigurdsson, H., Jonsdottir, T., Hrafinkelsson, J., Johannsson, J., Sveinsson, T., Myrdal, G., Grimsson, H. N., Jonsson, T., von Holst, S., Werelius, B., Margolin, S., Lindblom, A., Mayordomo, J. I., Haiman, C. A., Kiemeny, L. A., Johannsson, O. T., Gulcher, J. R., Thorsteinsdottir, U., Kong, A., & Stefansson, K. (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.. *Nature genetics* 39, 865--869.

23. Gudmundsson, J., Sulem, P., Rafnar, T., Bergthorsson, J. T., Manolescu, A., Gudbjartsson, D., Agnarsson, B. A., Sigurdsson, A., Benediktsdottir, K. R., Blondal, T., Jakobsdottir, M., Stacey, S. N., Kostic, J., Kristinsson, K. T., Birgisdottir, B., Ghosh, S., Magnusdottir, D. N., Thorlacius, S., Thorleifsson, G., Zheng, S. L., Sun, J., Chang, B.-L., Elmore, J. B., Breyer, J. P., McReynolds, K. M., Bradley, K. M., Yaspan, B. L., Wiklund, F., Stattin, P., Lindström, S., Adami, H.-O., McDonnell, S. K., Schaid, D. J., Cunningham, J. M., Wang, L., Cerhan, J. R., St Sauver, J. L., Isaacs, S. D., Wiley, K. E., Partin, A. W., Walsh, P. C., Polo, S., Ruiz-Echarri, M., Navarrete, S., Fuertes, F., Saez, B., Godino, J., Weijerman, P. C., Swinkels, D. W., Aben, K. K., Witjes, J. A., Suarez, B. K., Helfand, B. T., Frigge, M. L., Kristjansson, K., Ober, C., Jonsson, E., Einarsson, G. V., Xu, J., Gronberg, H., Smith, J. R., Thibodeau, S. N., Isaacs, W. B., Catalona, W. J., Mayordomo, J. I., Kiemeny, L. A., Barkardottir, R. B., Gulcher, J. R., Thorsteinsdottir, U., Kong, A., & Stefansson, K. (2008) Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer.. *Nature genetics* 40, 281--283.

24. Pritchard, J. K. (2001) Are rare variants responsible for susceptibility to complex diseases?. *American journal of human genetics* 69, 124--137.

25. Bodmer, W. & Tomlinson, I. (2010) Rare genetic variants and the risk of cancer.. *Current opinion in genetics & development* 20, 262--267.
26. Ewing, C. M., Ray, A. M., Lange, E. M., Zuhlke, K. A., Robbins, C. M., Tembe, W. D., Wiley, K. E., Isaacs, S. D., Johng, D., Wang, Y., Bizon, C., Yan, G., Gielzak, M., Partin, A. W., Shanmugam, V., Izatt, T., Sinari, S., Craig, D. W., Zheng, S. L., Walsh, P. C., Montie, J. E., Xu, J., Carpten, J. D., Isaacs, W. B., & Cooney, K. A. (2012) Germline mutations in HOXB13 and prostate-cancer risk.. *The New England journal of medicine* 366, 141--149.
27. Sodha, N., Mantoni, T. S., Tavgigian, S. V., Eeles, R., & Garrett, M. D. (2006) Rare germ line CHEK2 variants identified in breast cancer families encode proteins that show impaired activation.. *Cancer research* 66, 8966--8970.
28. Ponting, C. P. & Lunter, G. (2006) Signatures of adaptive evolution within human non-coding sequence.. *Human molecular genetics* 15 Spec No 2, R170--R175.
29. Ghildiyal, M. & Zamore, P. D. (2009) Small silencing RNAs: an expanding universe.. *Nature reviews. Genetics* 10, 94--108.
30. Kleinjan, D. A. & van Heyningen, V. (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease.. *American journal of human genetics* 76, 8--32.
31. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome.. *Nature* 489, 57--74.
32. modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., Lin, M. F., Washietl, S., Arshinoff, B. I., Ay, F., Meyer, P. E., Robine, N., Washington, N. L., Di Stefano, L., Berezikov, E., Brown, C. D., Candeias, R., Carlson, J. W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M. Y., Will, S., Alekseyenko, A. A., Artieri, C., Booth, B. W., Brooks, A. N., Dai, Q., Davis, C. A., Duff, M. O., Feng, X., Gorchakov, A. A., Gu, T., Henikoff, J. G., Kapranov, P., Li, R., MacAlpine, H. K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S. K., Riddle, N. C., Sakai, A., Samsonova, A., Sandler, J. E., Schwartz, Y. B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K. H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S. E., Brent, M. R., Cherbas, L., Elgin, S. C. R., Gingeras, T. R., Grossman, R., Hoskins, R. A., Kaufman, T. C., Kent, W., Kuroda, M. I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J. W., Ren, B., Russell, S., Cherbas, P., Graveley, B. R., Lewis, S., Micklem, G., Oliver, B., Park, P. J., Celniker, S. E., Henikoff, S., Karpen, G. H., Lai, E. C., MacAlpine, D. M., Stein, L. D., White, K. P., & Kellis, M. (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE.. *Science (New York, N.Y.)* 330, 1787--1797.
33. Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., Yip, K. Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., Alves, P., Chateigner, A., Perry, M., Morris, M., Auerbach, R. K., Feng, X., Leng, J., Vielle, A., Niu, W., Rhrissorrakrai, K., Agarwal, A., Alexander, R. P., Barber, G., Brdlik, C. M., Brennan, J., Brouillet, J. J., Carr, A., Cheung, M.-S., Clawson, H., Contrino, S., Dannenberg, L. O., Dernburg, A. F., Desai, A., Dick, L., Dosé, A. C., Du, J., Egelhofer, T., Ercan, S., Euskirchen, G., Ewing, B., Feingold, E. A., Gassmann, R., Good, P. J., Green, P., Gullier, F., Gutwein, M., Guyer, M. S., Habegger, L., Han, T., Henikoff, J. G., Henz, S. R., Hinrichs, A., Holster, H., Hyman, T., Iniguez, A. L., Jannette, J., Jensen, M., Kato, M., Kent, W. J., Kephart, E., Khivansara, V., Khurana, E., Kim, J. K., Kolasinska-Zwierz, P., Lai, E. C., Latorre, I., Leahey, A., Lewis, S., Lloyd, P., Lochovsky, L., Lowdon, R. F., Lubling, Y., Lyne, R., MacCoss, M., Mackowiak, S. D., Mangone, M., McKay, S., Mecnas, D., Merrihew, G., Miller, D. M., Muroyama, A., Murray, J. I., Ooi, S.-L., Pham, H., Phippen, T., Preston, E. A., Rajewsky, N., Rättsch, G., Rosenbaum, H., Rozowsky, J., Rutherford, K., Ruzanov, P., Sarov, M., Sasidharan, R., Sboner, A., Scheid, P., Segal, E., Shin, H., Shou, C., Slack, F. J., Slightam, C., Smith, R.,

- Spencer, W. C., Stinson, E. O., Taing, S., Takasaki, T., Vafeados, D., Voronina, K., Wang, G., Washington, N. L., Whittle, C. M., Wu, B., Yan, K.-K., Zeller, G., Zha, Z., Zhong, M., Zhou, X., modENCODE Consortium, Ahringer, J., Strome, S., Gunsalus, K. C., Micklem, G., Liu, X. S., Reinke, V., Kim, S. K., Hillier, L. W., Henikoff, S., Piano, F., Snyder, M., Stein, L., Lieb, J. D., & Waterston, R. H. (2010) Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project.. *Science (New York, N.Y.)* 330, 1775--1787.
34. Lappalainen, T., Sammeth, M., Friedländer, M. R., 't Hoen, P. A. C., Monlong, J., Rivas, M. A., González-Porta, M., Kurbatova, N., Griebel, T., Ferreira, P. G., Barann, M., Wieland, T., Greger, L., van Itersson, M., Almlöf, J., Ribeca, P., Pulyakhina, I., Esser, D., Giger, T., Tikhonov, A., Sultan, M., Bertier, G., MacArthur, D. G., Lek, M., Lizano, E., Buermans, H. P. J., Padioleau, I., Schwarzmayr, T., Karlberg, O., Ongen, H., Kilpinen, H., Beltran, S., Gut, M., Kahlem, K., Amstislavskiy, V., Stegle, O., Pirinen, M., Montgomery, S. B., Donnelly, P., McCarthy, M. I., Flicek, P., Strom, T. M., Geuvadis Consortium, Lehrach, H., Schreiber, S., Sudbrak, R., Carracedo, A., Antonarakis, S. E., Häslér, R., Syvänen, A.-C., van Ommen, G.-J., Brazma, A., Meitinger, T., Rosenstiel, P., Guigó, R., Gut, I. G., Estivill, X., & Dermitzakis, E. T. (2013) Transcriptome and genome sequencing uncovers functional variation in humans.. *Nature* 501, 506--511.
35. Montgomery, S. B., Sammeth, M., Gutierrez-Arcelus, M., Lach, R. P., Ingle, C., Nisbett, J., Guigo, R., & Dermitzakis, E. T. (2010) Transcriptome genetics using second generation sequencing in a Caucasian population.. *Nature* 464, 773--777.
36. Pickrell, J. K., Marioni, J. C., Pai, A. A., Degner, J. F., Engelhardt, B. E., Nkadori, E., Veyrieras, J.-B., Stephens, M., Gilad, Y., & Pritchard, J. K. (2010) Understanding mechanisms underlying human gene expression variation with RNA sequencing.. *Nature* 464, 768--772.
37. Battle, A., Mostafavi, S., Zhu, X., Potash, J. B., Weissman, M. M., McCormick, C., Haudenschild, C. D., Beckman, K. B., Shi, J., Mei, R., Urban, A. E., Montgomery, S. B., Levinson, D. F., & Koller, D. (2014) Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals.. *Genome research* 24, 14--24.
38. Gjoneska, E., Pfenning, A. R., Mathys, H., Quon, G., Kundaje, A., Tsai, L.-H., & Kellis, M. (2015) Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease.. *Nature* 518, 365--369.
39. Farh, K. K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., Shores, N., Whitton, H., Ryan, R. J. H., Shishkin, A. A., Hatan, M., Carrasco-Alfonso, M. J., Mayer, D., Luckey, C. J., Patsopoulos, N. A., De Jager, P. L., Kuchroo, V. K., Epstein, C. B., Daly, M. J., Hafler, D. A., & Bernstein, B. E. (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants.. *Nature* 518, 337--343.
40. Wijetunga, N. A., Delahaye, F., Zhao, Y. M., Golden, A., Mar, J. C., Einstein, F. H., & Grealley, J. M. (2014) The meta-epigenomic structure of purified human stem cell populations is defined at cis-regulatory sequences.. *Nature communications* 5, 5195.
41. Zhou, X., Li, D., Zhang, B., Lowdon, R. F., Rockweiler, N. B., Sears, R. L., Madden, P. A. F., Smirnov, I., Costello, J. F., & Wang, T. (2015) Epigenomic annotation of genetic variants using the Roadmap Epigenome Browser.. *Nature biotechnology* , .
42. Ernst, J. & Kellis, M. (2015) Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues.. *Nature biotechnology* 33, 364--376.
43. GTEx Consortium (2013) The Genotype-Tissue Expression (GTEx) project.. *Nature genetics* 45, 580--585.
44. Rivas, M. A., Pirinen, M., Conrad, D. F., Lek, M., Tsang, E. K., Karczewski, K. J., Maller, J. B.,

- Kukurba, K. R., DeLuca, D. S., Fromer, M., Ferreira, P. G., Smith, K. S., Zhang, R., Zhao, F., Banks, E., Poplin, R., Ruderfer, D. M., Purcell, S. M., Tukiainen, T., Minikel, E. V., Stenson, P. D., Cooper, D. N., Huang, K. H., Sullivan, T. J., Nedzel, J., GTEx Consortium, Geuvadis Consortium, Bustamante, C. D., Li, J. B., Daly, M. J., Guigo, R., Donnelly, P., Ardlie, K., Sammeth, M., Dermitzakis, E. T., McCarthy, M. I., Montgomery, S. B., Lappalainen, T., & MacArthur, D. G. (2015) Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome.. *Science (New York, N.Y.)* 348, 666--669.
45. Melé M1, Ferreira PG2, R. F. D. D. M. J. S. M. Y. T. G. J. P. D. S. T. J. R. S. A. D. S. N. A. G. C. W. F. L. T. C. M. G. G. D. E. A. K. G. R. (2015) Human genomics. The human transcriptome across tissues and individuals.. *Science (New York, N.Y.)* 348, 660-5.
46. GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans.. *Science (New York, N.Y.)* 348, 648--660.
47. Baran, Y., Subramaniam, M., Biton, A., Tukiainen, T., Tsang, E. K., Rivas, M. A., Pirinen, M., Gutierrez-Arcelus, M., Smith, K. S., Kukurba, K. R., Zhang, R., Eng, C., Torgerson, D. G., Urbanek, C., GTEx Consortium, Li, J. B., Rodriguez-Santana, J. R., Burchard, E. G., Seibold, M. A., MacArthur, D. G., Montgomery, S. B., Zaitlen, N. A., & Lappalainen, T. (2015) The landscape of genomic imprinting across diverse adult human tissues.. *Genome research* , .
48. Won, K.-J., Zhang, X., Wang, T., Ding, B., Raha, D., Snyder, M., Ren, B., & Wang, W. (2013) Comparative annotation of functional regions in the human genome using epigenomic data.. *Nucleic acids research* 41, 4423--4432.
49. Montgomery, S. B., Lappalainen, T., Gutierrez-Arcelus, M., & Dermitzakis, E. T. (2011) Rare and common regulatory variation in population-scale sequenced human genomes.. *PLoS genetics* 7, e1002144.
50. Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., & Fields, S. (2010) High-resolution mapping of protein sequence-function relationships.. *Nature methods* 7, 741--746.
51. Starita, L. M., Pruneda, J. N., Lo, R. S., Fowler, D. M., Kim, H. J., Hiatt, J. B., Shendure, J., Brzovic, P. S., Fields, S., & Klevit, R. E. (2013) Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis.. *Proceedings of the National Academy of Sciences of the United States of America* 110, E1263--E1272.
52. Araya, C. L., Fowler, D. M., Chen, W., Muniez, I., Kelly, J. W., & Fields, S. (2012) A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function.. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16858--16863.
53. Pitt, J. N. & Ferré-D'Amaré, A. R. (2010) Rapid construction of empirical RNA fitness landscapes.. *Science (New York, N.Y.)* 330, 376--379.
54. Rozowsky, J., Euskirchen, G., Auerbach, R. K., Zhang, Z. D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., & Gerstein, M. B. (2009) PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls.. *Nature biotechnology* 27, 66--75.
55. Cheng, C., Min, R., & Gerstein, M. (2011) TIP: a probabilistic method for identifying transcription factor target genes from ChIP-seq binding profiles.. *Bioinformatics (Oxford, England)* 27, 3221--3227.
56. Yip, K. Y. & Gerstein, M. (2009) Training set expansion: an approach to improving the reconstruction of biological networks from limited and uneven reliable interactions.. *Bioinformatics (Oxford, England)* 25, 243--250.
57. Yip, K. Y., Alexander, R. P., Yan, K.-K., & Gerstein, M. (2010) Improved reconstruction of in silico gene regulatory networks by integrating knockout and perturbation data.. *PloS one* 5, e8121.
58. Yip, K. Y., Cheng, C., Bhardwaj, N., Brown, J. B., Leng, J., Kundaje, A., Rozowsky, J., Birney, E., Bickel, P., Snyder, M., & Gerstein, M. (2012) Classification of human genomic regions based on

experimentally determined binding sites of more than 100 transcription-related factors.. *Genome biology* 13, R48.

59. Cheng, C., Shou, C., Yip, K. Y., & Gerstein, M. B. (2011) Genome-wide analysis of chromatin features identifies histone modification sensitive and insensitive yeast transcription factors.. *Genome biology* 12, R111.

60. Cheng, C., Yan, K.-K., Yip, K. Y., Rozowsky, J., Alexander, R., Shou, C., & Gerstein, M. (2011) A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets.. *Genome biology* 12, R15.

61. Cheng, C. & Gerstein, M. (2012) Modeling the relative relationship of transcription factor binding and histone modifications to gene expression levels in mouse embryonic stem cells.. *Nucleic acids research* 40, 553--568.

62. Cheng, C., Alexander, R., Min, R., Leng, J., Yip, K. Y., Rozowsky, J., Yan, K.-K., Dong, X., Djebali, S., Ruan, Y., Davis, C. A., Carninci, P., Lassman, T., Gingeras, T. R., Guigó, R., Birney, E., Weng, Z., Snyder, M., & Gerstein, M. (2012) Understanding transcriptional regulation by integrative analysis of transcription factor binding data.. *Genome research* 22, 1658--1667.

63. Gerstein, M. B., Kundaje, A., Hariharan, M., Landt, S. G., Yan, K.-K., Cheng, C., Mu, X. J., Khurana, E., Rozowsky, J., Alexander, R., Min, R., Alves, P., Abyzov, A., Addleman, N., Bhardwaj, N., Boyle, A. P., Cayting, P., Charos, A., Chen, D. Z., Cheng, Y., Clarke, D., Eastman, C., Euskirchen, G., Fietze, S., Fu, Y., Gertz, J., Grubert, F., Harmanci, A., Jain, P., Kasowski, M., Lacroute, P., Leng, J., Lian, J., Monahan, H., O'Geen, H., Ouyang, Z., Partridge, E. C., Patacsil, D., Pauli, F., Raha, D., Ramirez, L., Reddy, T. E., Reed, B., Shi, M., Slifer, T., Wang, J., Wu, L., Yang, X., Yip, K. Y., Zilberman-Schapira, G., Batzoglou, S., Sidow, A., Farnham, P. J., Myers, R. M., Weissman, S. M., & Snyder, M. (2012) Architecture of the human regulatory network derived from ENCODE data.. *Nature* 489, 91--100.

64. Nègre N1, Brown CD, M. L. B. C. M. S. W. U. K. P. E. M. L. P. S. R. L. Z. I. H. S. R. C. J. H. L. C. C. A. R. D. M. D. M. S. P. M. C. Z. J. S. S. S. L. V. A. B. N. G. A. H. D. M. D. M. M. V. K. B. H. W. R. G. M. R. S. G. R. R. B. P. J. K. M. W. K. (2011) A cis-regulatory map of the Drosophila genome.. *Nature* 471, 527-31.

65. Cheng, C., Yan, K.-K., Hwang, W., Qian, J., Bhardwaj, N., Rozowsky, J., Lu, Z. J., Niu, W., Alves, P., Kato, M., Snyder, M., & Gerstein, M. (2011) Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data.. *PLoS computational biology* 7, e1002190.

66. Yan, K.-K., Fang, G., Bhardwaj, N., Alexander, R. P., & Gerstein, M. (2010) Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks.. *Proceedings of the National Academy of Sciences of the United States of America* 107, 9186--9191.

67. Yu, H., Greenbaum, D., Xin Lu, H., Zhu, X., & Gerstein, M. (2004) Genomic analysis of essentiality within protein networks.. *Trends in genetics : TIG* 20, 227--231.

68. Yu, H., Zhu, X., Greenbaum, D., Karro, J., & Gerstein, M. (2004) TopNet: a tool for comparing biological sub-networks, correlating protein properties with topological statistics.. *Nucleic acids research* 32, 328--337.

69. Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., & Gerstein, M. (2007) The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics.. *PLoS computational biology* 3, e59.

70. Luscombe, N. M., Babu, M. M., Yu, H., Snyder, M., Teichmann, S. A., & Gerstein, M. (2004) Genomic analysis of regulatory network dynamics reveals large topological changes.. *Nature* 431, 308--312.

71. Gianoulis, T. A., Raes, J., Patel, P. V., Bjornson, R., Korbil, J. O., Letunic, I., Yamada, T., Paccanaro, A., Jensen, L. J., Snyder, M., Bork, P., & Gerstein, M. B. (2009) Quantifying environmental adaptation of

- metabolic pathways in metagenomics.. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1374--1379.
72. Yu, H., Paccanaro, A., Trifonov, V., & Gerstein, M. (2006) Predicting interactions in protein networks by completing defective cliques.. *Bioinformatics (Oxford, England)* 22, 823--829.
73. Kim, P. M., Korbelt, J. O., & Gerstein, M. B. (2007) Positive selection at the protein network periphery: evaluation in terms of structural constraints and cellular context.. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20274--20279.
74. Khurana, E., Fu, Y., Colonna, V., Mu, X. J., Kang, H. M., Lappalainen, T., Sboner, A., Lochovsky, L., Chen, J., Harmanci, A., Das, J., Abyzov, A., Balasubramanian, S., Beal, K., Chakravarty, D., Challis, D., Chen, Y., Clarke, D., Clarke, L., Cunningham, F., Evani, U. S., Flicek, P., Fragoza, R., Garrison, E., Gibbs, R., Gümüs, Z. H., Herrero, J., Kitabayashi, N., Kong, Y., Lage, K., Liliashvili, V., Lipkin, S. M., MacArthur, D. G., Marth, G., Muzny, D., Pers, T. H., Ritchie, G. R. S., Rosenfeld, J. A., Sisu, C., Wei, X., Wilson, M., Xue, Y., Yu, F., 1000 Genomes Project Consortium, Dermitzakis, E. T., Yu, H., Rubin, M. A., Tyler-Smith, C., & Gerstein, M. (2013) Integrative annotation of variants from 1092 humans: application to cancer genomics.. *Science (New York, N.Y.)* 342, 1235587.
75. Khurana, E., Fu, Y., Chen, J., & Gerstein, M. (2013) Interpretation of genomic variants using a unified biological network approach.. *PLoS computational biology* 9, e1002886.
76. Rozowsky, J., Abyzov, A., Wang, J., Alves, P., Raha, D., Harmanci, A., Leng, J., Bjornson, R., Kong, Y., Kitabayashi, N., Bhardwaj, N., Rubin, M., Snyder, M., & Gerstein, M. (2011) AlleleSeq: analysis of allele-specific expression and binding in a network framework.. *Molecular systems biology* 7, 522.
77. Shou, C., Bhardwaj, N., Lam, H. Y. K., Yan, K.-K., Kim, P. M., Snyder, M., & Gerstein, M. B. (2011) Measuring the evolutionary rewiring of biological networks.. *PLoS computational biology* 7, e1001050.
78. Boyle, A. P., Araya, C. L., Brdlik, C., Cayting, P., Cheng, C., Cheng, Y., Gardner, K., Hillier, L. W., Jannette, J., Jiang, L., Kasper, D., Kawli, T., Kheradpour, P., Kundaje, A., Li, J. J., Ma, L., Niu, W., Rehm, E. J., Rozowsky, J., Slatery, M., Spokony, R., Terrell, R., Vafeados, D., Wang, D., Weisdepp, P., Wu, Y.-C., Xie, D., Yan, K.-K., Feingold, E. A., Good, P. J., Pazin, M. J., Huang, H., Bickel, P. J., Brenner, S. E., Reinke, V., Waterston, R. H., Gerstein, M., White, K. P., Kellis, M., & Snyder, M. (2014) Comparative analysis of regulatory information and circuits across distant species.. *Nature* 512, 453--456.
79. Gerstein, M. B., Rozowsky, J., Yan, K.-K., Wang, D., Cheng, C., Brown, J. B., Davis, C. A., Hillier, L., Sisu, C., Li, J. J., Pei, B., Harmanci, A. O., Duff, M. O., Djebali, S., Alexander, R. P., Alver, B. H., Auerbach, R., Bell, K., Bickel, P. J., Boeck, M. E., Boley, N. P., Booth, B. W., Cherbas, L., Cherbas, P., Di, C., Dobin, A., Drenkow, J., Ewing, B., Fang, G., Fastuca, M., Feingold, E. A., Frankish, A., Gao, G., Good, P. J., Guigó, R., Hammonds, A., Harrow, J., Hoskins, R. A., Howald, C., Hu, L., Huang, H., Hubbard, T. J. P., Huynh, C., Jha, S., Kasper, D., Kato, M., Kaufman, T. C., Kitchen, R. R., Ladewig, E., Lagarde, J., Lai, E., Leng, J., Lu, Z., MacCoss, M., May, G., McWhirter, R., Merrihew, G., Miller, D. M., Mortazavi, A., Murad, R., Oliver, B., Olson, S., Park, P. J., Pazin, M. J., Perrimon, N., Pervouchine, D., Reinke, V., Reymond, A., Robinson, G., Samsonova, A., Saunders, G. I., Schlesinger, F., Sethi, A., Slack, F. J., Spencer, W. C., Stoiber, M. H., Strasbourger, P., Tanzer, A., Thompson, O. A., Wan, K. H., Wang, G., Wang, H., Watkins, K. L., Wen, J., Wen, K., Xue, C., Yang, L., Yip, K., Zaleski, C., Zhang, Y., Zheng, H., Brenner, S. E., Graveley, B. R., Celniker, S. E., Gingeras, T. R., & Waterston, R. (2014) Comparative analysis of the transcriptome across distant species.. *Nature* 512, 445--448.
80. Sisu, C., Pei, B., Leng, J., Frankish, A., Zhang, Y., Balasubramanian, S., Harte, R., Wang, D., Rutenberg-Schoenberg, M., Clark, W., Diekhans, M., Rozowsky, J., Hubbard, T., Harrow, J., & Gerstein, M. B. (2014) Comparative analysis of pseudogenes across three phyla.. *Proceedings of the National Academy of Sciences of the United States of America* 111, 13361--13366.

81. Yan, K.-K., Wang, D., Rozowsky, J., Zheng, H., Cheng, C., & Gerstein, M. (2014) OrthoClust: an orthology-based network framework for clustering data across multiple species.. *Genome biology* 15, R100.
82. Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., Alioto, T., Antoshechkin, I., Baer, M. T., Bar, N. S., Batut, P., Bell, K., Bell, I., Chakraborty, S., Chen, X., Chrast, J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M. J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Luo, O. J., Park, E., Persaud, K., Preall, J. B., Ribeca, P., Risk, B., Robyr, D., Sammeth, M., Schaffer, L., See, L.-H., Shahab, A., Skancke, J., Suzuki, A. M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S. E., Hannon, G., Giddings, M. C., Ruan, Y., Wold, B., Carninci, P., Guigó, R., & Gingeras, T. R. (2012) Landscape of transcription in human cells.. *Nature* 489, 101--108.
83. <http://brainspan.org> () Last accessed on 21st May 2015. , .
84. <http://exRNA.org> () Last accessed on 21st May 2015th January 2014. , .
85. Habegger, L., Sboner, A., Gianoulis, T. A., Rozowsky, J., Agarwal, A., Snyder, M., & Gerstein, M. (2011) RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries.. *Bioinformatics (Oxford, England)* 27, 281--283.
86. Du, J., Leng, J., Habegger, L., Sboner, A., McDermott, D., & Gerstein, M. (2012) IQSeq: integrated isoform quantification analysis based on next-generation sequencing.. *PloS one* 7, e29175.
87. Lu, Z. J., Yip, K. Y., Wang, G., Shou, C., Hillier, L. W., Khurana, E., Agarwal, A., Auerbach, R., Rozowsky, J., Cheng, C., Kato, M., Miller, D. M., Slack, F., Snyder, M., Waterston, R. H., Reinke, V., & Gerstein, M. B. (2011) Prediction and characterization of noncoding RNAs in *C. elegans* by integrating conservation, secondary structure, and high-throughput sequencing and array data.. *Genome research* 21, 276--285.
88. Mu, X. J., Lu, Z. J., Kong, Y., Lam, H. Y. K., & Gerstein, M. B. (2011) Analysis of genomic variation in non-coding elements using population-scale sequencing data from the 1000 Genomes Project.. *Nucleic acids research* 39, 7058--7076.
89. Pastinen, T. (2010) Genome-wide allele-specific analysis: insights into regulatory variation.. *Nature reviews. Genetics* 11, 533--538.
90. Birney, E., Lieb, J. D., Furey, T. S., Crawford, G. E., & Iyer, V. R. (2010) Allele-specific and heritable chromatin signatures in humans.. *Human molecular genetics* 19, R204--R209.
91. <http://alleleseq.gersteinlab.org> () Last accessed on 21st May 2015. , .
92. Garrison, E. & Marth, G. (2012) Haplotype-based variant detection from short-read sequencing. , .
93. Cooper, G. M., Stone, E. A., Asimenos, G., NISC Comparative Sequencing Program, Green, E. D., Batzoglou, S., & Sidow, A. (2005) Distribution and intensity of constraint in mammalian genomic sequence.. *Genome research* 15, 901--913.
94. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., Wu, Y.-C., Pfenning, A. R., Wang, X., Claussnitzer, M., Liu, Y., Coarfa, C., Harris, R. A., Shores, N., Epstein, C. B., Gjoneska, E., Leung, D., Xie, W., Hawkins, R. D., Lister, R., Hong, C., Gascard, P., Mungall, A. J., Moore, R., Chuah, E., Tam, A., Canfield, T. K., Hansen, R. S., Kaul, R., Sabo, P. J., Bansal, M. S., Carles, A., Dixon, J. R., Farh, K.-H., Feizi, S., Karlic, R., Kim, A.-R., Kulkarni, A., Li, D., Lowdon, R., Elliott, G., Mercer, T. R., Neph, S. J., Onuchic, V., Polak, P., Rajagopal, N., Ray, P., Sallari, R. C., Siebenthal, K. T., Sinnott-Armstrong, N.

- A., Stevens, M., Thurman, R. E., Wu, J., Zhang, B., Zhou, X., Beaudet, A. E., Boyer, L. A., De Jager, P. L., Farnham, P. J., Fisher, S. J., Haussler, D., Jones, S. J. M., Li, W., Marra, M. A., McManus, M. T., Sunyaev, S., Thomson, J. A., Tlsty, T. D., Tsai, L.-H., Wang, W., Waterland, R. A., Zhang, M. Q., Chadwick, L. H., Bernstein, B. E., Costello, J. F., Ecker, J. R., Hirst, M., Meissner, A., Milosavljevic, A., Ren, B., Stamatoyannopoulos, J. A., Wang, T., & Kellis, M. (2015) Integrative analysis of 111 reference human epigenomes.. *Nature* 518, 317--330.
95. Ziller, M. J., Edri, R., Yaffe, Y., Donaghey, J., Pop, R., Mallard, W., Issner, R., Gifford, C. A., Goren, A., Xing, J., Gu, H., Cacchiarelli, D., Tsankov, A. M., Epstein, C., Rinn, J. L., Mikkelsen, T. S., Kohlbacher, O., Gnirke, A., Bernstein, B. E., Elkabetz, Y., & Meissner, A. (2015) Dissecting neural differentiation regulatory networks through epigenetic footprinting.. *Nature* 518, 355--359.
96. Leung, D., Jung, I., Rajagopal, N., Schmitt, A., Selvaraj, S., Lee, A. Y., Yen, C.-A., Lin, S., Lin, Y., Qiu, Y., Xie, W., Yue, F., Hariharan, M., Ray, P., Kuan, S., Edsall, L., Yang, H., Chi, N. C., Zhang, M. Q., Ecker, J. R., & Ren, B. (2015) Integrative analysis of haplotype-resolved epigenomes across human tissues.. *Nature* 518, 350--354.
97. Blin, K., Dieterich, C., Wurmus, R., Rajewsky, N., Landthaler, M., & Akalin, A. (2015) DoRiNA 2.0--upgrading the doRiNA database of RNA interactions in post-transcriptional regulation.. *Nucleic acids research* 43, D160--D167.
98. Li, J.-H., Liu, S., Zhou, H., Qu, L.-H., & Yang, J.-H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data.. *Nucleic acids research* 42, D92--D97.
99. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., & Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.. *Cell* 141, 129--141.
100. Helwak, A., Kudla, G., Dudnakova, T., & Tollervey, D. (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding.. *Cell* 153, 654--665.
101. Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., & Bartel, D. P. (2011) Weak seed-pairing stability and high target-site abundance decrease the proficiency of *Isy-6* and other microRNAs.. *Nature structural & molecular biology* 18, 1139--1146.
102. Kheradpour, P., Ernst, J., Melnikov, A., Rogov, P., Wang, L., Zhang, X., Alston, J., Mikkelsen, T. S., & Kellis, M. (2013) Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay.. *Genome research* 23, 800--811.
103. Killela, P. J., Reitman, Z. J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L. A., Friedman, A. H., Friedman, H., Gallia, G. L., Giovanella, B. C., Grollman, A. P., He, T.-C., He, Y., Hruban, R. H., Jallo, G. I., Mandahl, N., Meeker, A. K., Mertens, F., Netto, G. J., Rasheed, B. A., Riggins, G. J., Rosenquist, T. A., Schiffman, M., Shih, I.-M., Theodorescu, D., Torbenson, M. S., Velculescu, V. E., Wang, T.-L., Wentzensen, N., Wood, L. D., Zhang, M., McLendon, R. E., Bigner, D. D., Kinzler, K. W., Vogelstein, B., Papadopoulos, N., & Yan, H. (2013) TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal.. *Proceedings of the National Academy of Sciences of the United States of America* 110, 6021--6026.
104. Vinagre, J., Almeida, A., Pópulo, H., Batista, R., Lyra, J., Pinto, V., Coelho, R., Celestino, R., Prazeres, H., Lima, L., Melo, M., da Rocha, A. G., Preto, A., Castro, P., Castro, L., Pardal, F., Lopes, J. M., Santos, L. L., Reis, R. M., Cameselle-Teijeiro, J., Sobrinho-Simões, M., Lima, J., Máximo, V., & Soares, P. (2013) Frequency of TERT promoter mutations in human cancers.. *Nature communications* 4, 2185.
105. Wang, Z. & Burge, C. B. (2008) Splicing regulation: from a parts list of regulatory elements to

- an integrated splicing code.. *RNA (New York, N.Y.)* 14, 802--813.
106. 1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R. E., Kang, H. M., Marth, G. T., & McVean, G. A. (2012) An integrated map of genetic variation from 1,092 human genomes.. *Nature* 491, 56--65.
107. Meyer, C. A. & Liu, X. S. (2014) Identifying and mitigating bias in next-generation sequencing methods for chromatin biology.. *Nature reviews. Genetics* 15, 709--721.
108. Srivastava, S. & Chen, L. (2010) A two-parameter generalized Poisson model to improve the analysis of RNA-seq data.. *Nucleic acids research* 38, e170.
109. Diaz, A., Park, K., Lim, D. A., & Song, J. S. (2012) Normalization, bias correction, and peak calling for ChIP-seq.. *Statistical applications in genetics and molecular biology* 11, Article 9.
110. DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernysky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data.. *Nature genetics* 43, 491--498.
111. Papaemmanuil, E., Cazzola, M., Boultwood, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J. S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., Godfrey, A. L., Rapado, I., Cvejic, A., Rance, R., McGee, C., Ellis, P., Mudie, L. J., Stephens, P. J., McLaren, S., Massie, C. E., Tarpey, P. S., Varela, I., Nik-Zainal, S., Davies, H. R., Shlien, A., Jones, D., Raine, K., Hinton, J., Butler, A. P., Teague, J. W., Baxter, E. J., Score, J., Galli, A., Della Porta, M. G., Travaglino, E., Groves, M., Tauro, S., Munshi, N. C., Anderson, K. C., El-Naggar, A., Fischer, A., Mustonen, V., Warren, A. J., Cross, N. C. P., Green, A. R., Futreal, P. A., Stratton, M. R., Campbell, P. J., & Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts.. *The New England journal of medicine* 365, 1384--1395.
112. Baca, S. C., Prandi, D., Lawrence, M. S., Mosquera, J. M., Romanel, A., Drier, Y., Park, K., Kitabayashi, N., MacDonald, T. Y., Ghandi, M., Van Allen, E., Kryukov, G. V., Sboner, A., Theurillat, J.-P., Soong, T. D., Nickerson, E., Auclair, D., Tewari, A., Beltran, H., Onofrio, R. C., Boysen, G., Guiducci, C., Barbieri, C. E., Cibulskis, K., Sivachenko, A., Carter, S. L., Saksena, G., Voet, D., Ramos, A. H., Winckler, W., Cipicchio, M., Ardlie, K., Kantoff, P. W., Berger, M. F., Gabriel, S. B., Golub, T. R., Meyerson, M., Lander, E. S., Elemento, O., Getz, G., Demichelis, F., Rubin, M. A., & Garraway, L. A. (2013) Punctuated evolution of prostate cancer genomes.. *Cell* 153, 666--677.
113. Berger, M. F., Lawrence, M. S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A. Y., Sboner, A., Esgueva, R., Pflueger, D., Sougnez, C., Onofrio, R., Carter, S. L., Park, K., Habegger, L., Ambrogio, L., Fennell, T., Parkin, M., Saksena, G., Voet, D., Ramos, A. H., Pugh, T. J., Wilkinson, J., Fisher, S., Winckler, W., Mahan, S., Ardlie, K., Baldwin, J., Simons, J. W., Kitabayashi, N., MacDonald, T. Y., Kantoff, P. W., Chin, L., Gabriel, S. B., Gerstein, M. B., Golub, T. R., Meyerson, M., Tewari, A., Lander, E. S., Getz, G., Rubin, M. A., & Garraway, L. A. (2011) The genomic complexity of primary human prostate cancer.. *Nature* 470, 214--220.
114. Greene, C. S., Krishnan, A., Wong, A. K., Ricciotti, E., Zelaya, R. A., Himmelstein, D. S., Zhang, R., Hartmann, B. M., Zaslavsky, E., Sealfon, S. C., Chasman, D. I., FitzGerald, G. A., Dolinski, K., Grosser, T., & Troyanskaya, O. G. (2015) Understanding multicellular function and disease with human tissue-specific networks.. *Nature genetics* , .
115. Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S. T., Abeyasinghe, S., Krawczak, M., & Cooper, D. N. (2003) Human Gene Mutation Database (HGMD): 2003 update.. *Human mutation* 21, 577--581.
116. Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott,

- D. R. (2014) ClinVar: public archive of relationships among sequence variation and human phenotype.. *Nucleic acids research* 42, D980--D985.
117. Kwasniewski, J. C., Mogno, I., Myers, C. A., Corbo, J. C., & Cohen, B. A. (2012) Complex effects of nucleotide variants in a mammalian cis-regulatory element.. *Proceedings of the National Academy of Sciences of the United States of America* 109, 19498--19503.
118. Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants.. *Nature genetics* 46, 310--315.
119. Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J. D., & Lin, X. (2013) Sequence kernel association tests for the combined effect of rare and common variants.. *American journal of human genetics* 92, 841--853.
120. ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/working/20130103_high_cov_trio_bams/ () Last accessed on 21st May 2015. , .
121. ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/complete_genomics_indices/ () Last accessed on 21st May 2015. , .
122. <http://www.uk10k.org/> () Last accessed on 21st May 2015. , .
123. Rausch, T., Jones, D. T. W., Zapatka, M., Stütz, A. M., Zichner, T., Weischenfeldt, J., Jäger, N., Remke, M., Shih, D., Northcott, P. A., Pfaff, E., Tica, J., Wang, Q., Massimi, L., Witt, H., Bender, S., Pleier, S., Cin, H., Hawkins, C., Beck, C., von Deimling, A., Hans, V., Brors, B., Eils, R., Scheurlen, W., Blake, J., Benes, V., Kulozik, A. E., Witt, O., Martin, D., Zhang, C., Porat, R., Merino, D. M., Wasserman, J., Jabado, N., Fontebasso, A., Bullinger, L., Rucker, F. G., Döhner, K., Döhner, H., Koster, J., Molenaar, J. J., Versteeg, R., Kool, M., Tabori, U., Malkin, D., Korshunov, A., Taylor, M. D., Lichter, P., Pfister, S. M., & Korbel, J. O. (2012) Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations.. *Cell* 148, 59--71.
124. Guda, K., Moinova, H., He, J., Jamison, O., Ravi, L., Natale, L., Lutterbaugh, J., Lawrence, E., Lewis, S., Willson, J. K. V., Lowe, J. B., Wiesner, G. L., Parmigiani, G., Barnholtz-Sloan, J., Dawson, D. W., Velculescu, V. E., Kinzler, K. W., Papadopoulos, N., Vogelstein, B., Willis, J., Gerken, T. A., & Markowitz, S. D. (2009) Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers.. *Proceedings of the National Academy of Sciences of the United States of America* 106, 12921--12925.
125. Fu, W., O'Connor, T. D., Jun, G., Kang, H. M., Abecasis, G., Leal, S. M., Gabriel, S., Rieder, M. J., Altshuler, D., Shendure, J., Nickerson, D. A., Bamshad, M. J., NHLBI Exome Sequencing Project, & Akey, J. M. (2013) Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants.. *Nature* 493, 216--220.
126. Vandenbroucke, I., Van Marck, H., Verhasselt, P., Thys, K., Mostmans, W., Dumont, S., Van Eygen, V., Coen, K., Tuefferd, M., & Aerssens, J. (2011) Minor variant detection in amplicons using 454 massive parallel pyrosequencing: experiences and considerations for successful applications.. *BioTechniques* 51, 167--177.
127. Chakravarty, D., Sboner, A., Nair, S. S., Giannopoulou, E., Li, R., Hennig, S., Mosquera, J. M., Pauwels, J., Park, K., Kossai, M., MacDonald, T. Y., Fontugne, J., Erho, N., Vergara, I. A., Ghadessi, M., Davicioni, E., Jenkins, R. B., Palanisamy, N., Chen, Z., Nakagawa, S., Hirose, T., Bander, N. H., Beltran, H., Fox, A. H., Elemento, O., & Rubin, M. A. (2014) The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer.. *Nature communications* 5, 5383.
128. Roy, S., Chakravarty, D., Cortez, V., De Mukhopadhyay, K., Bandyopadhyay, A., Ahn, J.-M., Raj, G. V., Tekmal, R. R., Sun, L., & Vadlamudi, R. K. (2012) Significance of PELP1 in ER-negative breast

cancer metastasis.. *Molecular cancer research : MCR* 10, 25--33.

129. Dimple, C., Nair, S. S., Rajhans, R., Pitcheswara, P. R., Liu, J., Balasenthil, S., Le, X.-F., Burow, M. E., Auersperg, N., Tekmal, R. R., Broaddus, R. R., & Vadlamudi, R. K. (2008) Role of PELP1/MNAR signaling in ovarian tumorigenesis.. *Cancer research* 68, 4902--4909.
130. Lamesch, P., Li, N., Milstein, S., Fan, C., Hao, T., Szabo, G., Hu, Z., Venkatesan, K., Bethel, G., Martin, P., Rogers, J., Lawlor, S., McLaren, S., Dricot, A., Borick, H., Cusick, M. E., Vandenhaute, J., Dunham, I., Hill, D. E., & Vidal, M. (2007) hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes.. *Genomics* 89, 307--315.
131. (last accessed on may 22nd 2015) <http://www.yulab.org/Supp/MutPrimer>. , .
132. Bartsch, G., Horninger, W., Klocker, H., Pelzer, A., Bektic, J., Oberaigner, W., Schennach, H., Schäfer, G., Frauscher, F., Boniol, M., Severi, G., Robertson, C., Boyle, P., & Tyrol Prostate Cancer Screening Group (2008) Tyrol Prostate Cancer Demonstration Project: early detection, treatment, outcome, incidence and mortality.. *BJU international* 101, 809--816.
133. Oberaigner, W., Horninger, W., Klocker, H., Schönitzer, D., Stühlinger, W., & Bartsch, G. (2006) Reduction of prostate cancer mortality in Tyrol, Austria, after introduction of prostate-specific antigen testing.. *American journal of epidemiology* 164, 376--384.
134. <http://edrn.nci.nih.gov/> () Last accessed on 21st May 2015. , .
135. Garritano, S., Romanel, A., Ciribilli, Y., Bisio, A., Gavoci, A., Inga, A., & Demichelis, F. (2015) In-silico identification and functional validation of allele-dependent AR enhancers.. *Oncotarget* 6, 4816--4828.
136. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system.. *Nature protocols* 8, 2281--2308.
137. Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., & Zhang, F. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases.. *Nature biotechnology* 31, 827--832.
138. Beltran (2015) . *JAMA oncology* in press. , .
139. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013) RNA-programmed genome editing in human cells.. *eLife* 2, e00471.
140. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems.. *Science (New York, N.Y.)* 339, 819--823.
141. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013) RNA-guided human genome engineering via Cas9.. *Science (New York, N.Y.)* 339, 823--826.
142. Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.. *Cell* 153, 910--918.