In silico and wet lab approaches to study transcriptional regulation

Hestand, M.S.

Abbreviations

Commonly used Abbreviations in this thesis:

bp = base pair(s)
CAGE = Cap Analysis of Gene Expression
ChIP = chromatin immuniprecipitation
ChIP-(on-)chip = ChIP hybridized to a microarray
ChIP-PET = ChIP with paired-end ditag sequencing
ChIP-seq = ChIP with next-generation sequencing
DeepCAGE = next-generation sequencing of CAGE sequences
DeepSAGE = next-generation sequencing of SAGE sequences
HAT = histone acetyltransferase
Kb = kilo base(s)
NGS = next-generation sequencers or sequencing
PCR = polymerase chain reaction
PWM = position weight matrix
qPCR = quantitative (real-time) PCR
RT-PCR = reverse transcriptase PCR
SAGE = Serial Analysis of Gene Expression
TF = transcription factor
TFBS = transcription factor binding site
TSS = transcription start site
UTR = untranslated region

The four nucleobases of DNA:

A = adenine
T = thymine
C = cytosine
G = guanine

Bibliography

[1] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de laBastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K.,

Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., deJong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001) *Nature* **409(6822)**, 860–921.

[2] Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D. M., Eddy, S. R., Elnitski, L., Emes, R. D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T. A., Green, E. D., Gregory, S., Guigo, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W. J., Kirby, A., Kolbe, D. L., Korf, I., Kucherlapati, R. S., Kulbokas, E. J., Kulp, D., Landers, T., Leger, J. P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D. R., Mardis, E. R., Matthews, L., Mauceli, E., Mayer, J. H., McCarthy, M., McCombie, W. R., McLaren, S., McLay, K., McPherson, J. D., Meldrim, J., Meredith, B., Mesirov, J. P., Miller, W., Miner, T. L., Mongin, E., Montgomery, K. T., Morgan, M., Mott, R., Mullikin, J. C., Muzny, D. M., Nash, W. E., Nelson, J. O., Nhan, M. N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M. J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K. H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C. S., Poliakov, A., Ponce, T. C., Ponting, C. P., Potter, S., Quail, M., Reymond, A., Roe, B. A., Roskin, K. M., Rubin, E. M., Rust, A. G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M. S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J. B., Slater, G., Smit, A., Smith, D. R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J. P., Niederhausern, A. C. V., Wade, C. M., Wall, M., Weber, R. J., Weiss, R. B., Wendl, M. C., West, A. P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R. K., Winter, E., Worley,

K. C., Wyman, D., Yang, S., Yang, S.-P., Zdobnov, E. M., Zody, M. C., and Lander, E. S. (2002) *Nature* **420(6915)**, 520–62.

[3] Hubbard, T. J. P., Aken, B. L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., Coates, G., Fairley, S., Fitzgerald, S., Fernandez-Banet, J., Gordon, L., Graf, S., Haider, S., Hammond, M., Holland, R., Howe, K., Jenkinson, A., Johnson, N., Kahari, A., Keefe, D., Keenan, S., Kinsella, R., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Rios, D., Schuster, M., Slater, G., Smedley, D., Spooner, W., Spudich, G., Trevanion, S., Vilella, A., Vogel, J., White, S., Wilder, S., Zadissa, A., Birney, E., Cunningham, F., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Kasprzyk, A., Proctor, G., Smith, J., Searle, S., and Flicek, P. (2009) *Nucleic Acids Res* **37(Database issue)**, D690–7.

[4] Kasprzyk, A., Keefe, D., Smedley, D., London, D., Spooner, W., Melsopp, C., Hammond, M., Rocca-Serra, P., Cox, T., and Birney, E. (2004) *Genome Res* **14(1)**, 160–9.

[5] Lannigan, D. A. and Notides, A. C. (1989) *Proc Natl Acad Sci U S A* **86(3)**, 863–7.

[6] Lee, J. S., Lee, C. H., and Chung, J. H. (1998) *Proc Natl Acad Sci U S A* **95(3)**, 969–74.

[7] Reed, B. D., Charos, A. E., Szekely, A. M., Weissman, S. M., and Snyder, M. (2008) *PLoS Genet* **4(7)**, e1000133.

[8] Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A. M., Taylor, M. S., Engstrom, P. G., Frith, M. C., Forrest, A. R. R., Alkema, W. B., Tan, S. L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S. M., Wells, C. A., Orlando, V., Wahlestedt, C., Liu, E. T., Harbers, M., Kawai, J., Bajic, V. B., Hume, D. A., and Hayashizaki, Y. (2006) *Nat Genet* **38(6)**, 626–35.

[9] Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., and Ren, B. (2005) *Nature* **436(7052)**, 876–80.

[10] Gardiner-Garden, M. and Frommer, M. (1987) *J Mol Biol* **196(2)**, 261–82.

[11] Wei, C.-L., Wu, Q., Vega, V. B., Chiu, K. P., Ng, P., Zhang, T., Shahab, A., Yong, H. C., Fu, Y., Weng, Z., Liu, J., Zhao, X. D., Chew, J.-L., Lee, Y. L., Kuznetsov, V. A., Sung, W.-K., Miller, L. D., Lim, B., Liu, E. T., Yu, Q., Ng, H.-H., and Ruan, Y. (2006) *Cell* **124(1)**, 207–19.

[12] Arney, K. L. and Fisher, A. G. (2004) *J Cell Sci* **117(Pt 19)**, 4355–63.

[13] Miele, A. and Dekker, J. (2008) *Mol Biosyst* **4(11)**, 1046–57.

[14] King, M. C. and Wilson, A. C. (1975) *Science* **188(4184)**, 107–16.

[15] Satyanarayana, A. and Kaldis, P. (2009) *Oncogene* **28(33)**, 2925–39.

[16] Malumbres, M. and Barbacid, M. (2009) *Nat Rev Cancer* **9(3)**, 153–66.

[17] Leake, R. (1996) *Ann N Y Acad Sci* **784**, 252–62.

[18] Polager, S. and Ginsberg, D. (2009) *Nat Rev Cancer* **9(10)**, 738–48.

[19] Grossman, S. R. (2001) *Eur J Biochem* **268(10)**, 2773–8.

[20] Yao, T. P., Oh, S. P., Fuchs, M., Zhou, N. D., Ch'ng, L. E., Newsome, D., Bronson, R. T., Li, E., Livingston, D. M., and Eckner, R. (1998) *Cell* **93(3)**, 361–72.

[21] Tanaka, Y., Naruse, I., Hongo, T., Xu, M., Nakahata, T., Maekawa, T., and Ishii, S. (2000) *Mech Dev* **95(1-2)**, 133–45.

[22] Charge, S. B. P. and Rudnicki, M. A. (2004) *Physiol Rev* **84(1)**, 209–38.

[23] Asakura, A. and Rudnicki, M. (2002) *In: Mouse Development* , 253278.

[24] Pownall, M. E., Gustafsson, M. K., and Emerson, C. P. J. (2002) *Annu Rev Cell Dev Biol* **18**, 747–83.

[25] Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y., and Dynlacht, B. D. (2005) *Genes Dev* **19(5)**, 553–69.

[26] Martin, P. T. (2003) *Curr Opin Pharmacol* **3(3)**, 300–8.

[27] Chomczynski, P. and Sacchi, N. (2006) *Nat Protoc* **1(2)**, 581–5.

[28] Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) *Science* **270(5235)**, 484–7.

[29] Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T., Fukuda, S., Sasaki, D., Podhajska, A., Harbers, M., Kawai, J., Carninci, P., and Hayashizaki, Y. (2003) *Proc Natl Acad Sci U S A* **100(26)**, 15776–81.

[30] Nielsen, K. L., Hogh, A. L., and Emmersen, J. (2006) *Nucleic Acids Res* **34(19)**, e133.

[31] Valen, E., Pascarella, G., Chalk, A., Maeda, N., Kojima, M., Kawazu, C., Murata, M., Nishiyori, H., Lazarevic, D., Motti, D., Marstrand, T. T., Tang, M.-H. E., Zhao, X., Krogh, A., Winther, O., Arakawa, T., Kawai, J., Wells, C., Daub, C., Harbers, M., Hayashizaki, Y., Gustincich, S., Sandelin, A., and Carninci, P. (2009) *Genome Res* **19(2)**, 255–65.

[32] Harbers, M. and Carninci, P. (2005) *Nat Methods* **2(7)**, 495–502.

[33] Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., and Surani, M. A. (2009) *Nat Methods* **6(5)**, 377–82.

[34] Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T. L., Wilson, C. J., Bell, S. P., and Young, R. A. (2000) *Science* **290(5500)**, 2306–9.

[35] Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O. L., He, A., Marra, M., Snyder, M., and Jones, S. (2007) *Nat Methods* **4(8)**, 651–7.

[36] 'tHoen, P. A. C., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., deMenezes, R. X., Boer, J. M., vanOmmen, G.-J. B., and denDunnen, J. T. (2008) *Nucleic Acids Res* **36(21)**, e141.

[37] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J Mol Biol* **215(3)**, 403–10.

[38] Kent, W. J. (2002) *Genome Res* **12(4)**, 656–64.

[39] Li, H., Ruan, J., and Durbin, R. (2008) *Genome Res* **18(11)**, 1851–8.

[40] Smith, A. D., Xuan, Z., and Zhang, M. Q. (2008) *BMC Bioinformatics* **9**, 128.

[41] Schatz, M. C. (2009) *Bioinformatics* **25(11)**, 1363–9.

[42] Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009) *Genome Biol* **10(3)**, R25.

[43] Li, H. and Durbin, R. (2009) *Bioinformatics* **25(14)**, 1754–60.

[44] Zerbino, D. R. and Birney, E. (2008) *Genome Res* **18(5)**, 821–9.

[45] Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006) *Nucleic Acids Res* **34(Web Server issue)**, W369–73.

[46] Bailey, T. L. and Elkan, C. (1994) *Proc Int Conf Intell Syst Mol Biol* **2**, 28–36.

[47] Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F., and Wootton, J. C. (1993) *Science* **262(5131)**, 208–14.

[48] Neuwald, A. F., Liu, J. S., and Lawrence, C. E. (1995) *Protein Sci* **4(8)**, 1618–32.

[49] Pavesi, G., Mauri, G., and Pesole, G. (2004) *Brief Bioinform* **5(3)**, 217–36.

[50] Stormo, G. D. (1990) *Methods Enzymol* **183**, 211–21.

[51] Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A. E., Kel-Margoulis, O. V., Kloos, D.-U., Land, S., Lewicki-Potapov, B., Michael, H., Munch, R., Reuter, I., Rotert, S., Saxel, H., Scheer, M., Thiele, S., and Wingender, E. (2003) *Nucleic Acids Res* **31(1)**, 374–8.

[52] TRANSFAC website: *www.biobase-international.com/cgi-bin/biobase/transfac/start.cgi.*

[53] Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W. W., and Lenhard, B. (2004) *Nucleic Acids Res* **32(Database issue)**, D91–4.

[54] JASPAR website:
http://jaspar.genereg.net/.

[55] Kel, A. E., Gossling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., and Wingender, E. (2003) *Nucleic Acids Res* **31(13)**, 3576–9.

[56] Hoffman, M. M. and Birney, E. (2010) *Genome Res* **20(5)**, 685–92.

[57] Elkon, R., Linhart, C., Sharan, R., Shamir, R., and Shiloh, Y. (2003) *Genome Res* **13(5)**, 773–80.

[58] Gumucio, D. L., Shelton, D. A., Zhu, W., Millinoff, D., Gray, T., Bock, J. H., Slightom, J. L., and Goodman, M. (1996) *Mol Phylogenet Evol* **5(1)**, 18–32.

[59] Hardison, R. C., Oeltjen, J., and Miller, W. (1997) *Genome Res* **7(10)**, 959–66.

[60] Sui, S. J. H., Fulton, D. L., Arenillas, D. J., Kwon, A. T., and Wasserman, W. W. (2007) *Nucleic Acids Res* **35(Web Server issue)**, W245–52.

[61] Hooghe, B., Hulpiau, P., vanRoy, F., and Bleser, P. D. (2008) *Nucleic Acids Res* **36(Web Server issue)**, W128–32.

[62] Hubbard, T. J. P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Fitzgerald, S., Fernandez-Banet, J., Graf, S., Haider, S., Hammond, M., Herrero, J., Holland, R., Howe, K., Howe, K., Johnson, N., Kahari, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Melsopp, C., Megy, K., Meidl, P., Ouverdin, B., Parker, A., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Severin, J., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A., Vogel, J., White, S., Wood, M., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Flicek, P., Kasprzyk, A., Proctor, G., Searle, S., Smith, J., Ureta-Vidal, A., and Birney, E. (2007) *Nucleic Acids Res* **35(Database issue)**, D610–7.

[63] sorttable.js: *http://www.kryogenix.org/code/browser/sorttable/*.

[64] Kobes, R. to access the cephes math library, by Moshier, SL. math::cephes perl interface:
http://search.cpan.org/dist/Math-Cephes/lib/Math/Cephes.pod.

[65] Schwartz, S., Kent, W. J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R. C., Haussler, D., and Miller, W. (2003) *Genome Res* **13(1)**, 103–7.

[66] Cao, Y., Kumar, R. M., Penn, B. H., Berkes, C. A., Kooperberg, C., Boyer, L. A., Young, R. A., and Tapscott, S. J. (2006) *EMBO J* **25(3)**, 502–11.

[67] Smyth, G., Yang, Y., and Speed, T. (2003) *Methods Mol Biol* **224**, 111–136.

[68] Gentleman, R., Carey, V., Dudoit, S., Irizarry, R., and Huber, W. (2005) Bioinformatics and Computational Biology Solutions using R and Bioconductor., Springer, New York.

[69] Alibes, A., Yankilevich, P., Canada, A., and Diaz-Uriarte, R. (2007) *BMC Bioinformatics* **8**, 9.

[70] Benjamini, Y. and Hochberg, Y. (1995) *J R Statist Soc B* **57**, 289–300.

[71] Hubbard, T., Andrews, D., Caccamo, M., Cameron, G., Chen, Y., Clamp, M., Clarke, L., Coates, G., Cox, T., Cunningham, F., Curwen, V., Cutts, T., Down, T., Durbin, R., Fernandez-Suarez, X. M., Gilbert, J., Hammond, M., Herrero, J., Hotz, H., Howe, K., Iyer, V., Jekosch, K., Kahari, A., Kasprzyk, A., Keefe, D., Keenan, S., Kokocinsci, F., London, D., Longden, I., McVicker, G., Melsopp, C., Meidl, P., Potter, S., Proctor, G., Rae, M., Rios, D., Schuster, M., Searle, S., Severin, J., Slater, G., Smedley, D., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Storey, R., Trevanion, S., Ureta-Vidal, A., Vogel, J., White, S., Woodwark, C., and Birney, E. (2005) *Nucleic Acids Res* **33(Database issue)**, D447–53.

[72] Edmondson, D. G., Brennan, T. J., and Olson, E. N. (1991) *J Biol Chem* **266(32)**, 21343–6.

[73] Banerjee-Basu, S. and Buonanno, A. (1993) *Mol Cell Biol* **13(11)**, 7019–28.

[74] Brunetti, A. and Goldfine, I. D. (1990) *J Biol Chem* **265(11)**, 5960–3.

[75] Wyzykowski, J. C., Winata, T. I., Mitin, N., Taparowsky, E. J., and Konieczny, S. F. (2002) *Mol Cell Biol* **22(17)**, 6199–208.

[76] Hestand, M. S., vanGalen, M., Villerius, M. P., vanOmmen, G.-J. B., denDunnen, J. T., and 'tHoen, P. A. C. (2008) *BMC Bioinformatics* **9**, 495.

[77] Tokovenko, B., Golda, R., Protas, O., Obolenskaya, M., and El'skaya, A. (2009) *Nucleic Acids Res* **37(7)**, e49.

[78] Zambelli, F., Pesole, G., and Pavesi, G. (2009) *Nucleic Acids Res* **37(Web Server issue)**, W247–52.

[79] Marstrand, T. T., Frellsen, J., Moltke, I., Thiim, M., Valen, E., Retelska, D., and Krogh, A. (2008) *PLoS One* **3(2)**, e1623.

[80] Zheng, J., Wu, J., and Sun, Z. (2003) *Nucleic Acids Res* **31(7)**, 1995–2005.

[81] Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., and Gaul, U. (2008) *Nature* **451(7178)**, 535–40.

[82] Sinha, S. (2006) *Bioinformatics* **22(14)**, e454–63.

[83] Durbin, R., Eddy, S., Krogh, A., and Mitchison, G. (1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, Cambridge, UK.

[84] Roider, H. G., Manke, T., O'Keeffe, S., Vingron, M., and Haas, S. A. (2009) *Bioinformatics* **25(4)**, 435–42.

[85] Roider, H. G., Lenhard, B., Kanhere, A., Haas, S. A., and Vingron, M. (2009) *Nucleic Acids Res* **37(19)**, 6305–15.

[86] Ng, P. and Keich, U. (2008) *Genome Inform* **21**, 15–26.

[87] Ng, P. and Keich, U. (2008) *Bioinformatics* **24(19)**, 2256–7.

[88] Khatri, P. and Draghici, S. (2005) *Bioinformatics* **21(18)**, 3587–95.

[89] Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) *Nucleic Acids Res* **37(1)**, 1–13.

[90] Jelier, R., Schuemie, M. J., Veldhoven, A., Dorssers, L. C. J., Jenster, G., and Kors, J. A. (2008) *Genome Biol* **9(6)**, R96.

[91] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) *Nat Genet* **25(1)**, 25–9.

[92] Camon, E., Barrell, D., Lee, V., Dimmer, E., and Apweiler, R. (2004) *In Silico Biol* **4(1)**, 5–6.

[93] Ramos, Y. F., Hestand, M. S., Verlaan, M., Krabbendam, E., Ariyurek, Y., vanGalen, M., vanDam, H., vanOmmen, G. J., denDunnen, J. T., Zantema, A., and 'tHoen, P. A. (2010) *Nucleic Acids Res* **Epub**, 2010 Apr 30.

[94] Rice, P., Longden, I., and Bleasby, A. (2000) *Trends Genet* **16(6)**, 276–7.

[95] Slater, G. S. C. and Birney, E. (2005) *BMC Bioinformatics* **6**, 31.

[96] Jothi, R., Cuddapah, S., Barski, A., Cui, K., and Zhao, K. (2008) *Nucleic Acids Res* **36(16)**, 5221–31.

[97] Valouev, A., Johnson, D. S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R. M., and Sidow, A. (2008) *Nat Methods* **5(9)**, 829–34.

[98] Kharchenko, P. V., Tolstorukov, M. Y., and Park, P. J. (2008) *Nat Biotechnol* **26(12)**, 1351–9.

[99] Fejes, A. P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S. J. M. (2008) *Bioinformatics* **24(15)**, 1729–30.

[100] Jongeneel, C. V., Iseli, C., Stevenson, B. J., Riggins, G. J., Lal, A., Mackay, A., Harris, R. A., O'Hare, M. J., Neville, A. M., Simpson, A. J. G., and Strausberg, R. L. (2003) *Proc Natl Acad Sci U S A* **100(8)**, 4702–5.

[101] Quail, M. A., Kozarewa, I., Smith, F., Scally, A., Stephens, P. J., Durbin, R., Swerdlow, H., and Turner, D. J. (2008) *Nat Methods* **5(12)**, 1005–10.

[102] Dohm, J. C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008) *Nucleic Acids Res* **36(16)**, e105.

[103] Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., and Haussler, D. (2002) *Genome Res* **12(6)**, 996–1006.

[104] Farnham, P. J. (2009) *Nat Rev Genet* **10(9)**, 605–16.

[105] Moore, M. J. and Proudfoot, N. J. (2009) *Cell* **136(4)**, 688–700.

[106] Sonenberg, N. and Hinnebusch, A. G. (2009) *Cell* **136(4)**, 731–45.

[107] Lu, T.-Y., Kao, C.-F., Lin, C.-T., Huang, D.-Y., Chiu, C.-Y., Huang, Y.-S., and Wu, H.-C. (2009) *J Cell Biochem* **108(1)**, 315–25.

[108] Selaru, F. M., David, S., Meltzer, S. J., and Hamilton, J. P. (2009) *Am J Gastroenterol* **104(8)**, 1910–2.

[109] Szyf, M. (2009) *Clin Rev Allergy Immunol* , Epub ahead of print.

[110] Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) *Cell* **129(4)**, 823–37.

[111] Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008) *Nat Genet* **40(7)**, 897–903.

[112] Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S., and Bernstein, B. E. (2007) *Nature* **448(7153)**, 553–60.

[113] Welboren, W.-J., vanDriel, M. A., Janssen-Megens, E. M., vanHeeringen, S. J., Sweep, F. C., Span, P. N., and Stunnenberg, H. G. (2009) *EMBO J* **28(10)**, 1418–28.

[114] Smeenk, L., vanHeeringen, S. J., Koeppel, M., vanDriel, M. A., Bartels, S. J. J., Akkers, R. C., Denissov, S., Stunnenberg, H. G., and Lohrum, M. (2008) *Nucleic Acids Res* **36(11)**, 3639–54.

[115] Wederell, E. D., Bilenky, M., Cullum, R., Thiessen, N., Dagpinar, M., Delaney, A., Varhol, R., Zhao, Y., Zeng, T., Bernier, B., Ingham, M., Hirst, M., Robertson, G., Marra, M. A., Jones, S., and Hoodless, P. A. (2008) *Nucleic Acids Res* **36(14)**, 4549–64.

[116] Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007) *Science* **316(5830)**, 1497–502.

[117] Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W., and Zhao, K. (2009) *Cell* **138(5)**, 1019–31.

[118] Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E. M., and Pennacchio, L. A. (2009) *Nature* **457(7231)**, 854–8.

[119] Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Calcar, S. V., Qu, C., Ching, K. A., Wang, W., Weng, Z., Green, R. D., Crawford, G. E., and Ren, B. (2007) *Nat Genet* **39(3)**, 311–8.

[120] Kalkhoven, E. (2004) *Biochem Pharmacol* **68(6)**, 1145–55.

[121] Goodman, R. H. and Smolik, S. (2000) *Genes Dev* **14(13)**, 1553–77.

[122] Roelfsema, J. H. and Peters, D. J. M. (2007) *Expert Rev Mol Med* **9(23)**, 1–16.

[123] Iyer, N. G., Ozdag, H., and Caldas, C. (2004) *Oncogene* **23(24)**, 4225–31.

[124] Chrivia, J. C., Kwok, R. P., Lamb, N., Hagiwara, M., Montminy, M. R., and Goodman, R. H. (1993) *Nature* **365(6449)**, 855–9.

[125] Duyndam, M. C., vanDam, H., Smits, P. H., Verlaan, M., van derEb, A. J., and Zantema, A. (1999) *Oncogene* **18(14)**, 2311–21.

[126] Puri, P. L., Avantaggiati, M. L., Balsano, C., Sang, N., Graessmann, A., Giordano, A., and Levrero, M. (1997) *EMBO J* **16(2)**, 369–83.

[127] Partanen, A., Motoyama, J., and Hui, C. C. (1999) *Int J Dev Biol* **43(6)**, 487–94.

[128] Kung, A. L., Rebel, V. I., Bronson, R. T., Ch'ng, L. E., Sieff, C. A., Livingston, D. M., and Yao, T. P. (2000) *Genes Dev* **14(3)**, 272–7.

[129] Shikama, N., Lutz, W., Kretzschmar, R., Sauter, N., Roth, J.-F., Marino, S., Wittwer, J., Scheidweiler, A., and Eckner, R. (2003) *EMBO J* **22(19)**, 5175–85.

[130] Bordoli, L., Husser, S., Luthi, U., Netsch, M., Osmani, H., and Eckner, R. (2001) *Nucleic Acids Res* **29(21)**, 4462–71.

[131] Tullai, J. W., Schaffer, M. E., Mullenbrock, S., Sholder, G., Kasif, S., and Cooper, G. M. (2007) *J Biol Chem* **282(33)**, 23981–95.

[132] Denissov, S., vanDriel, M., Voit, R., Hekkelman, M., Hulsen, T., Hernandez, N., Grummt, I., Wehrens, R., and Stunnenberg, H. (2007) *EMBO J* **26(4)**, 944–54.

[133] Freeman, M. F. and Tukey, J. W. (1950) *Ann Math Statist* **21(4)**, 607–611.

[134] Dennis, G. J., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) *Genome Biol* **4(5)**, P3.

[135] Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Ostell, J., Pruitt, K. D., Schuler, G. D., Shumway, M., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T. A., Wagner, L., and Yaschenko, E. (2008) *Nucleic Acids Res* **36(Database issue)**, D13–21.

[136] Deng, L., de laFuente, C., Fu, P., Wang, L., Donnelly, R., Wade, J. D., Lambert, P., Li, H., Lee, C. G., and Kashanchi, F. (2000) *Virology* **277(2)**, 278–95.

[137] Chow, C.-W. and Davis, R. J. (2006) *Cell* **127(5)**, 887–890.

[138] Gilchrist, D. A., Fargo, D. C., and Adelman, K. (2009) *Methods* **48(4)**, 398–408.

[139] Cho, H., Orphanides, G., Sun, X., Yang, X. J., Ogryzko, V., Lees, E., Nakatani, Y., and Reinberg, D. (1998) *Mol Cell Biol* **18(9)**, 5355–63.

[140] Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., and Wahlestedt, C. (2005) *Science* **309(5740)**, 1564–6.

[141] Gordon, S., Akopyan, G., Garban, H., and Bonavida, B. (2006) *Oncogene* **25(8)**, 1125–42.

[142] Seto, E., Lewis, B., and Shenk, T. (1993) *Nature* **365(6445)**, 462–4.

[143] Zhou, Q., Gedrich, R. W., and Engel, D. A. (1995) *J Virol* **69(7)**, 4323–30.

[144] Morin, R., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T., McDonald, H., Varhol, R., Jones, S., and Marra, M. (2008) *Biotechniques* **45(1)**, 81–94.

[145] Sterrenburg, E., Turk, R., 'tHoen, P. A. C., vanDeutekom, J. C. T., Boer, J. M., vanOmmen, G.-J. B., and denDunnen, J. T. (2004) *Neuromuscul Disord* **14(8-9)**, 507–18.

[146] Turk, R., Sterrenburg, E., van derWees, C. G. C., deMeijer, E. J., deMenezes, R. X., Groh, S., Campbell, K. P., Noguchi, S., vanOmmen, G. J. B., denDunnen, J. T., and 'tHoen, P. A. C. (2006) *FASEB J* **20(1)**, 127–9.

[147] Jelier, R., 'tHoen, P. A. C., Sterrenburg, E., denDunnen, J. T., vanOmmen, G.-J. B., Kors, J. A., and Mons, B. (2008) *BMC Bioinformatics* **9**, 291.

[148] Tomczak, K. K., Marinescu, V. D., Ramoni, M. F., Sanoudou, D., Montanaro, F., Han, M., Kunkel, L. M., Kohane, I. S., and Beggs, A. H. (2004) *FASEB J* **18(2)**, 403–5.

[149] Sartorelli, V. and Caretti, G. (2005) *Curr Opin Genet Dev* **15(5)**, 528–35.

[150] Vencio, R. Z. N., Brentani, H., Patrao, D. F. C., and Pereira, C. A. B. (2004) *BMC Bioinformatics* **5**, 119.

[151] Huber, W., vonHeydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) *Bioinformatics* **18 Suppl 1**, S96–104.

[152] Cohen, C. D., Klingenhoff, A., Boucherot, A., Nitsche, A., Henger, A., Brunner, B., Schmid, H., Merkle, M., Saleem, M. A., Koller, K.-P., Werner, T., Grone, H.-J., Nelson, P. J., and Kretzler, M. (2006) *Proc Natl Acad Sci U S A* **103(15)**, 5682–7.

[153] Dohr, S., Klingenhoff, A., Maier, H., deAngelis, M. H., Werner, T., and Schneider, R. (2005) *Nucleic Acids Res* **33(3)**, 864–72.

[154] Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I. F., Soboleva, A., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Muertter, R. N., and Edgar, R. (2009) *Nucleic Acids Res* **37(Database issue)**, D885–90.

[155] Velculescu, V. E., Madden, S. L., Zhang, L., Lash, A. E., Yu, J., Rago, C., Lal, A., Wang, C. J., Beaudry, G. A., Ciriello, K. M., Cook, B. P., Dufault, M. R., Ferguson, A. T., Gao, Y., He, T. C., Hermeking, H., Hiraldo, S. K., Hwang, P. M., Lopez, M. A., Luderer, H. F., Mathews, B., Petroziello, J. M., Polyak, K., Zawel, L., and Kinzler, K. W. (1999) *Nat Genet* **23(4)**, 387–8.

[156] Buskin, J. N. and Hauschka, S. D. (1989) *Mol Cell Biol* **9(6)**, 2627–40.

[157] Olson, E. N. (1990) *Genes Dev* **4(9)**, 1454–61.

[158] Nishida, W., Nakamura, M., Mori, S., Takahashi, M., Ohkawa, Y., Tadokoro, S., Yoshida, K., Hiwada, K., Hayashi, K., and Sobue, K. (2002) *J Biol Chem* **277(9)**, 7308–17.

[159] Suzuki, H., Forrest, A. R. R., vanNimwegen, E., Daub, C. O., Balwierz, P. J., Irvine, K. M., Lassmann, T., Ravasi, T., Hasegawa, Y., deHoon, M. J. L., Katayama, S., Schroder, K., Carninci, P., Tomaru, Y., Kanamori-Katayama, M., Kubosaki, A., Akalin, A., Ando, Y., Arner, E., Asada, M., Asahara, H., Bailey, T., Bajic, V. B., Bauer, D., Beckhouse, A. G., Bertin, N., Bjorkegren, J., Brombacher, F., Bulger, E., Chalk, A. M., Chiba, J., Cloonan, N., Dawe, A., Dostie, J., Engstrom, P. G., Essack, M., Faulkner, G. J., Fink, J. L., Fredman, D., Fujimori, K., Furuno, M., Gojobori, T., Gough, J., Grimmond, S. M., Gustafsson, M., Hashimoto, M., Hashimoto, T., Hatakeyama, M., Heinzel, S., Hide, W., Hofmann, O., Hornquist, M., Huminiecki, L., Ikeo, K., Imamoto, N., Inoue, S., Inoue, Y., Ishihara, R., Iwayanagi, T., Jacobsen, A., Kaur, M., Kawaji, H., Kerr, M. C., Kimura, R., Kimura, S., Kimura, Y., Kitano, H., Koga, H., Kojima, T., Kondo, S., Konno, T., Krogh, A., Kruger, A., Kumar, A., Lenhard, B., Lennartsson, A., Lindow, M., Lizio, M., Macpherson, C., Maeda, N., Maher, C. A., Maqungo, M., Mar, J., Matigian, N. A., Matsuda, H., Mattick, J. S., Meier, S., Miyamoto, S., Miyamoto-Sato, E., Nakabayashi, K., Nakachi, Y., Nakano, M., Nygaard, S., Okayama, T., Okazaki, Y., Okuda-Yabukami, H., Orlando, V., Otomo, J., Pachkov, M., Petrovsky, N., Plessy, C., Quackenbush, J., Radovanovic, A., Rehli, M., Saito, R., Sandelin, A., Schmeier, S., Schonbach, C., Schwartz, A. S., Semple, C. A., Sera, M., Severin, J., Shirahige, K., Simons, C., Laurent, G. S., Suzuki, M., Suzuki, T., Sweet, M. J., Taft, R. J., Takeda, S., Takenaka, Y., Tan, K., Taylor, M. S., Teasdale, R. D., Tegner, J., Teichmann, S., Valen, E., Wahlestedt, C., Waki, K., Waterhouse, A., Wells, C. A., Winther,

O., Wu, L., Yamaguchi, K., Yanagawa, H., Yasuda, J., Zavolan, M., Hume, D. A., Arakawa, T., Fukuda, S., Imamura, K., Kai, C., Kaiho, A., Kawashima, T., Kawazu, C., Kitazume, Y., Kojima, M., Miura, H., Murakami, K., Murata, M., Ninomiya, N., Nishiyori, H., Noma, S., Ogawa, C., Sano, T., Simon, C., Tagami, M., Takahashi, Y., Kawai, J., and Hayashizaki, Y. (2009) *Nat Genet* **41(5)**, 553–62.

[160] Fejes-Toth, Sotirova, Sachidanandam, Assaf, Hannon, Kapranov, Foissac, Willingham, Duttagupta, Dumais, and Gingeras (2009) *Nature* **457(7232)**, 1028–32.

[161] Sabourin, L. A. and Rudnicki, M. A. (2000) *Clin Genet* **57(1)**, 16–25.

[162] Funk, W. D. and Wright, W. E. (1992) *Proc Natl Acad Sci U S A* **89(20)**, 9484–8.

[163] Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., and Teichmann, S. A. (2004) *Curr Opin Struct Biol* **14(3)**, 283–91.

[164] Guigo, R. and Fickett, J. W. (1995) *J Mol Biol* **253(1)**, 51–60.

[165] Ozsolak, F., Platt, A. R., Jones, D. R., Reifenberger, J. G., Sass, L. E., McInerney, P., Thompson, J. F., Bowers, J., Jarosz, M., and Milos, P. M. (2009) *Nature* **461(7265)**, 814–8.

[166] Pushkarev, D., Neff, N. F., and Quake, S. R. (2009) *Nat Biotechnol* **27(9)**, 847–52.

[167] Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J. A., Andrews, R. M., Flicek, P., Boyle, P. J., Cao, H., Carter, N. P., Clelland, G. K., Davis, S., Day, N., Dhami, P., Dillon, S. C., Dorschner, M. O., Fiegler, H., Giresi, P. G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K. D., Johnson, B. E., Johnson, E. M., Frum, T. T., Rosenzweig, E. R., Karnani, N., Lee, K., Lefebvre, G. C., Navas, P. A., Neri, F., Parker, S. C. J., Sabo, P. J., Sandstrom, R., Shafer, A., Vetrie, D., Weaver, M., Wilcox, S., Yu, M., Collins, F. S., Dekker, J., Lieb, J. D., Tullius, T. D., Crawford, G. E., Sunyaev, S., Noble, W. S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I. L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H. A., Sekinger, E. A., Lagarde, J., Abril, J. F., Shahab, A., Flamm, C., Fried, C., Hackermuller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J. S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M. C., Thomas, D. J., Weirauch, M. T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K. G., Sung, W.-K., Ooi, H. S., Chiu, K. P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M. L., Valencia, A., Choo, S. W., Choo, C. Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T. G., Brown, J. B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C. N., Kai, C.,

Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J. S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R. M., Rogers, J., Stadler, P. F., Lowe, T. M., Wei, C.-L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S. E., Fu, Y., Green, E. D., Karaoz, U., Siepel, A., Taylor, J., Liefer, L. A., Wetterstrand, K. A., Good, P. J., Feingold, E. A., Guyer, M. S., Cooper, G. M., Asimenos, G., Dewey, C. N., Hou, M., Nikolaev, S., Montoya-Burgos, J. I., Loytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N. R., Holmes, I., Mullikin, J. C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W. J., Stone, E. A., Batzoglou, S., Goldman, N., Hardison, R. C., Haussler, D., Miller, W., Sidow, A., Trinklein, N. D., Zhang, Z. D., Barrera, L., Stuart, R., King, D. C., Ameur, A., Enroth, S., Bieda, M. C., Kim, J., Bhinge, A. A., Jiang, N., Liu, J., Yao, F., Vega, V. B., Lee, C. W. H., Ng, P., Shahab, A., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M. J., Inman, D., Singer, M. A., Richmond, T. A., Munn, K. J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J. C., Couttet, P., Bruce, A. W., Dovey, O. M., Ellis, P. D., Langford, C. F., Nix, D. A., Euskirchen, G., Hartman, S., Urban, A. E., Kraus, P., Calcar, S. V., Heintzman, N., Kim, T. H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C. K., Rosenfeld, M. G., Aldred, S. F., Cooper, S. J., Halees, A., Lin, J. M., Shulha, H. P., Zhang, X., Xu, M., Haidar, J. N. S., Yu, Y., Ruan, Y., Iyer, V. R., Green, R. D., Wadelius, C., Farnham, P. J., Ren, B., Harte, R. A., Hinrichs, A. S., Trumbower, H., Clawson, H., Hillman-Jackson, J., Zweig, A. S., Smith, K., Thakkapallayil, A., Barber, G., Kuhn, R. M., Karolchik, D., Armengol, L., Bird, C. P., deBakker, P. I. W., Kern, A. D., Lopez-Bigas, N., Martin, J. D., Stranger, B. E., Woodroffe, A., Davydov, E., Dimas, A., Eyras, E., Hallgrimsdottir, I. B., Huppert, J., Zody, M. C., Abecasis, G. R., Estivill, X., Bouffard, G. G., Guan, X., Hansen, N. F., Idol, J. R., Maduro, V. V. B., Maskeri, B., McDowell, J. C., Park, M., Thomas, P. J., Young, A. C., Blakesley, R. W., Muzny, D. M., Sodergren, E., Wheeler, D. A., Worley, K. C., Jiang, H., Weinstock, G. M., Gibbs, R. A., Graves, T., Fulton, R., Mardis, E. R., Wilson, R. K., Clamp, M., Cuff, J., Gnerre, S., Jaffe, D. B., Chang, J. L., Lindblad-Toh, K., Lander, E. S., Koriabine, M., Nefedov, M., Osoegawa, K., Yoshinaga, Y., Zhu, B., and deJong, P. J. (2007) *Nature* **447(7146)**, 799–816.

[168] Allo, M., Buggiano, V., Fededa, J. P., Petrillo, E., Schor, I., de laMata, M., Agirre, E., Plass, M., Eyras, E., Elela, S. A., Klinck, R., Chabot, B., and Kornblihtt, A. R. (2009) *Nat Struct Mol Biol* **16(7)**, 717–24.

[169] Yelin, R., Dahary, D., Sorek, R., Levanon, E. Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., Nemzer, S., Pinner, E., Walach, S., Bernstein, J., Savitsky, K., and Rotman, G. (2003) *Nat Biotechnol* **21(4)**, 379–86.

List of Publications

-Ramos YFM*, Hestand MS*, Verlaan M, Krabbendam E, Ariyurek Y, van Galen M, van Dam H, van Ommen GJ, den Dunnen JT, Zantema A*, 't Hoen PA*. Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. Nucleic Acids Res. 2010 Apr 30. [Epub ahead of print]
*=equal first or last authorship

-Hestand MS, van Galen M, Villerius MP, van Ommen GJ, den Dunnen JT, 't Hoen PA. 2008. CORE_TF: a user-friendly interface to identify evolutionary conserved transcription factor binding sites in sets of co-regulated genes. BMC Bioinformatics. Nov 26;9(1):495.

-Wiersma AC, Millon LV, Hestand MS, Van Oost BA, Bannasch DL. 2005. Canine COL4A3 and COL4A4: sequencing, mapping and genomic organization. DNA Seq. Aug;16(4):241-51.

-van den Berg L, Kwant L, Hestand MS, van Oost BA, Leegwater PA. 2005. Structure and variation of three canine genes involved in serotonin binding and transport: the serotonin receptor 1A gene (htr1A), serotonin receptor 2A gene (htr2A), and serotonin transporter gene (slc6A4). J Hered. 96(7):786-96.

Curriculum Vitae

Matthew Hestand was born January 3, 1979, in Lexington, KY, USA. He graduated from Tates Creek High School in Lexington, KY, in 1997. He then did a Bachelors of Science in Biology and Bachelors of Arts in Philosophy at the University of Kentucky, USA, graduating in December 2001. He was also a graduate of the university's Honors Program. During this time he worked and did research in the university's traditional equine blood-typing lab, including a mouse/alpaca syntenic mapping project. After his bachelors he worked as a lab technician in a virology lab at the Livestock Disease Diagnostic Center in Lexington, Kentucky.

From Fall 2002 until September 2004 he did a Masters in Chemistry with the Genomics and Bioinformatics program at the University of Utrecht, NL. This included several projects: sequencing and characterizing the canine serotonin transporter gene *SLC6A4*, (in silico) comparative genomics of the *MURR1* gene, and a masters thesis comparing EBI (Ensembl) and NCBI gene annotation and visualization.

In April 2005 he began a Center for Biomedical Genetics sponsored PhD at Leiden University, situated in the department of Human and Clinical Genetics at the Leiden University Medical Center under the supervision of Peter-Bram 't Hoen, Johan den Dunnen, and Gert-Jan van Ommen. The PhD focused on transcriptional regulation, with the results presented in this thesis. During this time he won the best application showcase at the NBIC (Netherlands BioInformatics Centre) Conference 2009 for his web-based application CORE_TF. Also, during and since this time he has been actively involved in education with supervision of several short and one long-term student projects and teaching and coordination of several next-generation sequencing courses.

Starting April 2009 he became the coordinator for next-generation sequencing bioinformatics at the Leiden Genome Technology Center in Leiden, NL as part of the NBIC BioAssist program.