

Supplement

37. R. E. Thurman *et al.*, *Nature* **489**, 75–82 (2012).
38. S. John *et al.*, *Nature Genetics* **43**, 264–268 (2011).
39. J. R. Hesselberth *et al.*, *Nature Methods* **6**, 283–289 (2009).
40. B. Langmead, C. Trapnell, M. Pop, S. Salzberg, *Genome Biology* **10**, R25+ (2009).
41. S. Neph *et al.*, *Bioinformatics* **28**, 1919–20 (2012).
42. K. D. Pruitt *et al.*, *Genome research* **19**, 1316–23 (2009).
43. F. Krueger, S. R. Andrews, *Bioinformatics (Oxford, England)* **27**, 1571–2 (2011).
44. P. Kumar, S. Henikoff, P. C. Ng, *Nature protocols* **4**, 1073–81 (2009).
45. I. A. Adzhubei *et al.*, *Nature methods* **7**, 248–9 (2010).
46. W. McLaren *et al.*, *Bioinformatics (Oxford, England)* **26**, 2069–70 (2010).
47. D. M. Altshuler *et al.*, *Nature* **467**, 52–8 (2010).

48. M. T. Maurano *et al.*, *Science* **337**, 1190–5 (2012).
49. F. Capon *et al.*, *Human molecular genetics* **17**, 1938–45 (2008).
50. P. E. Stuart *et al.*, *Nature genetics* **42**, 1000–4 (2010).
51. R. M. Durbin *et al.*, *Nature* **467**, 1061–1073 (2010).
52. V. Matys *et al.*, *Nucleic Acids Research* **34**, D108–D110 (2006).
53. J. C. Bryne *et al.*, *Nucleic Acids Research* **36**, D102–D106 (2008).
54. D. E. Newburger, M. L. Bulyk, *Nucleic Acids Research* **37**, D77–D82 (2009).
55. T. L. Bailey *et al.*, *Nucleic Acids Research* **37**, W202–W208 (2009).
56. G. E. Crooks, G. Hon, J.-M. M. Chandonia, S. E. Brenner, *Genome Research* **14**, 1188–1190 (2004).
57. T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, *Nature methods* **8**, 785–6 (2011).
58. A. Krogh, B. Larsson, G. von Heijne, E. L. Sonnhammer, *Journal of molecular biology* **305**, 567–80 (2001).
59. T. la Cour *et al.*, *Protein engineering, design & selection : PEDS* **17**, 527–36 (2004).
60. J. Gough, K. Karplus, R. Hughey, C. Chothia, *Journal of molecular biology* **313**, 903–19 (2001).
61. K. S. Pollard, M. J. Hubisz, K. R. Rosenbloom, A. Siepel, *Genome research* **20**, 110–21 (2010).
62. R. C. Griffiths, S. Tavaré, *Commun. Stat. Stoch. Models* **14**, 273–295 (1998).
63. J. A. Tennessen *et al.*, *Science* **337**, 64–9 (2012).