

## Extended Data Figure 1: Annotated sequence contigs from the MSY of five species.

All sequence features and BACs drawn to scale. **a–e**, Marmoset MSY. **f–j**, Mouse MSY. **k–o**, Rat MSY. **p–t**, Bull MSY. **u–y**, Opossum MSY. **a, f, k, p, u**, Schematic representation of assembled contigs and sequence classes: X-degenerate (yellow); ampliconic (blue); pseudoautosomal (green); heterochromatic (pink). Gaps shown in white. **b, g, l, q, v**, Positions of all intact, actively transcribed genes. Plus (+) strand above, minus (–) strand below. **c, h, m, r, w**, G + C content (%) calculated in a 100-kb sliding window with 1-kb steps. **d, i, n, s, x**, Alu (red), LINE (green), and endogenous retrovirus (blue) content (%) calculated in a 200-kb sliding window with 1-kb steps. **e, j, o, t, y**, Sequenced MSY BACs. Each bar represents the size and position of one BAC clone, labelled with the library identifier. **e**, BAC clones with no prefix are from the CHORI-259 library; BAC clones with “A” prefix are from the MARMAEX (Amplicon Express) library. **j**, All BAC clones are from the RPCI-24 library. **o**, BAC clones without prefix are from the RNAEX library; BAC clones with “E” prefix are from the RNECO library (both from Amplicon Express). **t**, BAC clones without prefix are from the CHORI-240 library; BAC clones with “E” prefix are from the BTDAEX library (both from Amplicon Express). **y**, BAC clones with no prefix are from the VMRC6 library; BAC clones with “A” prefix are from the MDAEX (Amplicon Express) library.