(A) Hnf4a

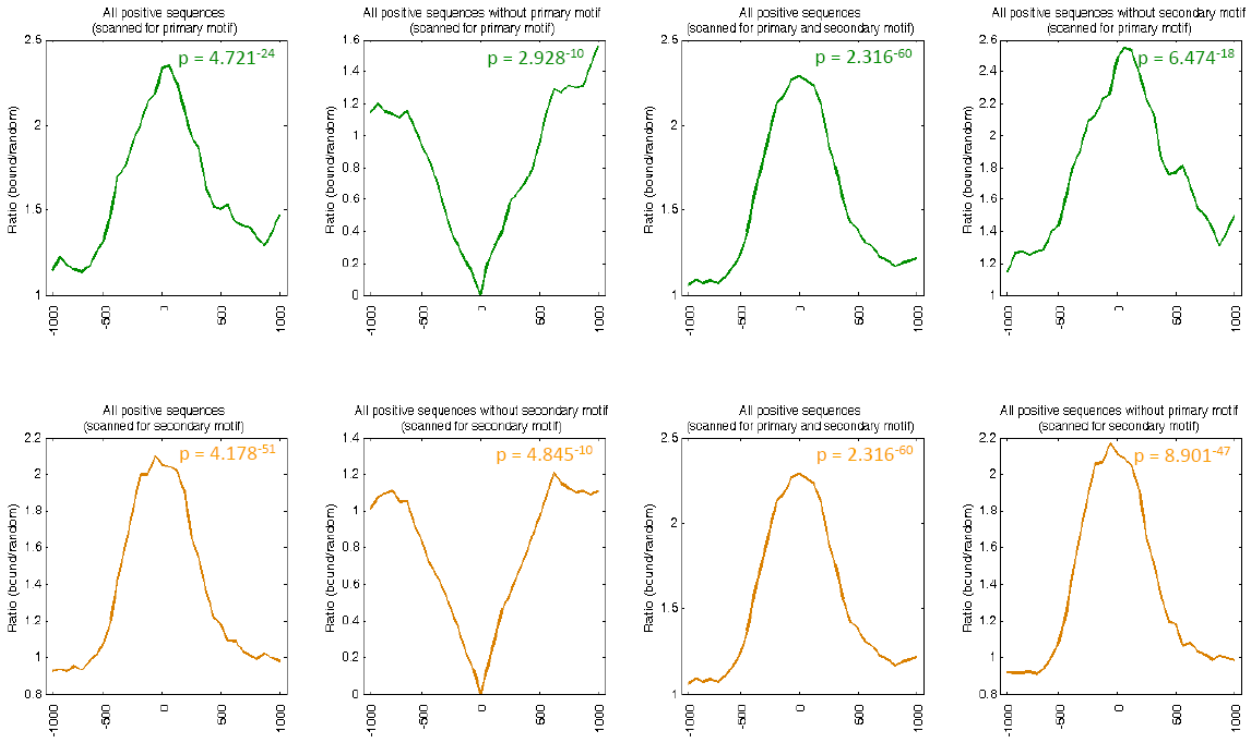
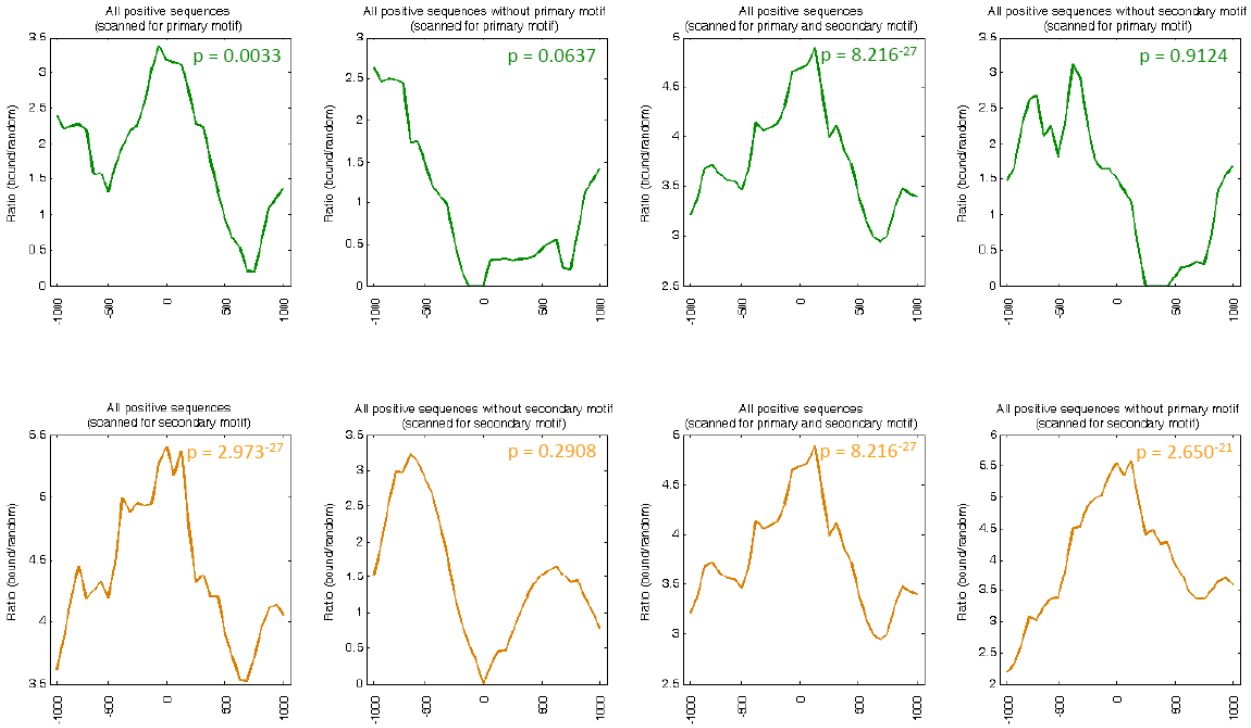
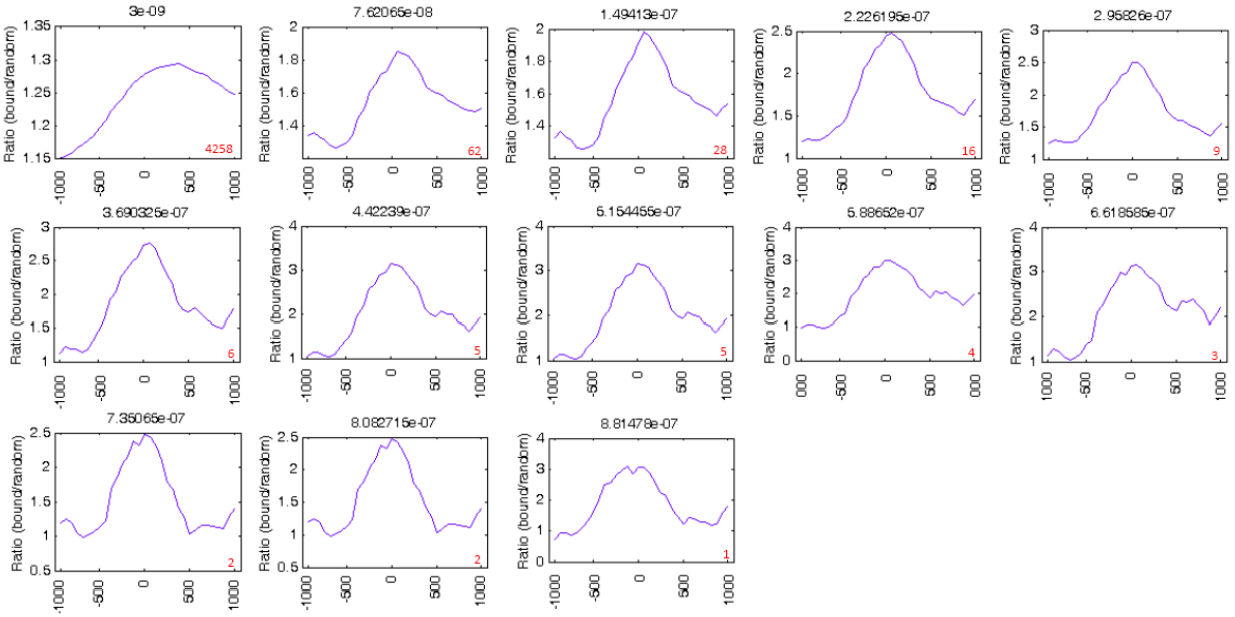


Figure S14: Enrichment of primary versus secondary motif 8-mers bound *in vitro* within genomic regions bound *in vivo*. Relative enrichment of *k*-mers corresponding to the primary versus secondary Seed-and-Wobble motifs within bound genomic regions in ChIP-chip data as compared to randomly selected sequences was calculated (see **Materials and Methods**) for **(A, C, D)** Hnf4a (Neilsen *et al.*, submitted; GEO accession #GSE7745) and **(B, E, F) (next page)** Bcl6b (*34*) (GEO accession #GSE7673). ChIP-chip 'bound' regions were identified according to the criteria of the respective studies (*34*)(Neilsen *et al.*, submitted). A window size of 500 bp with a step size of 100 bp was used. Either all 'bound' regions (far left, upper and lower rows), 'bound' regions lacking primary motif *k*-mers (second from left, upper row; far right, lower row) or 'bound' regions lacking secondary motif *k*-mers (far right, upper row; second from left, lower row) were considered for matches to primary motif *k*-mers (far left, second from left, and far right in upper row), secondary motif *k*-mers (far left, second from left, and far right in lower row), or either primary or secondary motif *k*-mers (second from right, upper and lower rows). The coarseness of the Bcl6 distributions is due to a smaller sample size of ChIP-chip 'bound' regions. The GOMER thresholds used in **(A)** are 2.958×10^{-7} and 8.419×10^{-7}, corresponding to 9 primary and 20 secondary 8-mers scanned, respectively for Hnf4a. The GOMER thresholds used for the data shown in **(B)** correspond to 1.513×10^{-6} and 3.294×10^{-7} corresponding to 4 primary and 17 secondary 8-mers scanned, respectively, for Bcl6b. *P*-values for enrichment of 8-mers within the bound genomic regions shown in each panel were calculated for the interval −250 to +250 by the Wilcoxon-Mann-Whitney rank sum test, comparing the number of occurrences per sequence in the bound set versus the background set. Enrichment plots at varying GOMER score thresholds (indicated above each plot in panels **C-F, next pages**) are shown in **(C, D)** for Hnf4a and **(E, F)** for Bcl6b for primary **(C, E)** versus secondary **(D, F)** motifs using a window size of 500 bp and a step size of 50 bp. Enrichment is generally observed across varying GOMER thresholds, with the exception that at permissive GOMER thresholds enrichment can be lost. Number of *k*-mers included at each GOMER threshold is indicated in red on each plot in panels **C-F**.

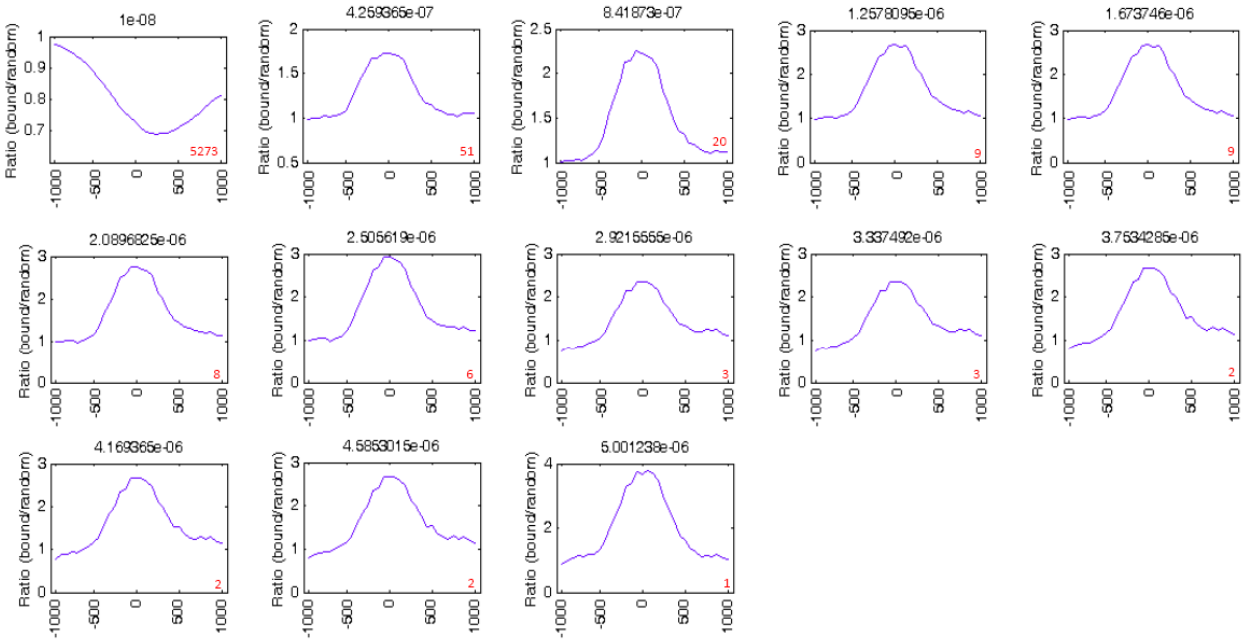
(B) Bcl6b


All positive sequences
(scanned for primary motif)
p = 0.0033
All positive sequences without primary motif
(scanned for primary motif)
p = 0.0637
All positive sequences
(scanned for primary and secondary motif)
p = 8.216-27
All positive sequences without secondary motif
(scanned for primary motif)
p = 0.9124
All positive sequences
(scanned for secondary motif)
p = 2.973-27
All positive sequences without secondary motif
(scanned for secondary motif)
p = 0.2908
All positive sequences
(scanned for primary and secondary motif)
p = 8.216-27
All positive sequences without primary motif
(scanned for secondary motif)
p = 2.650-21
Ratio (bound/random)

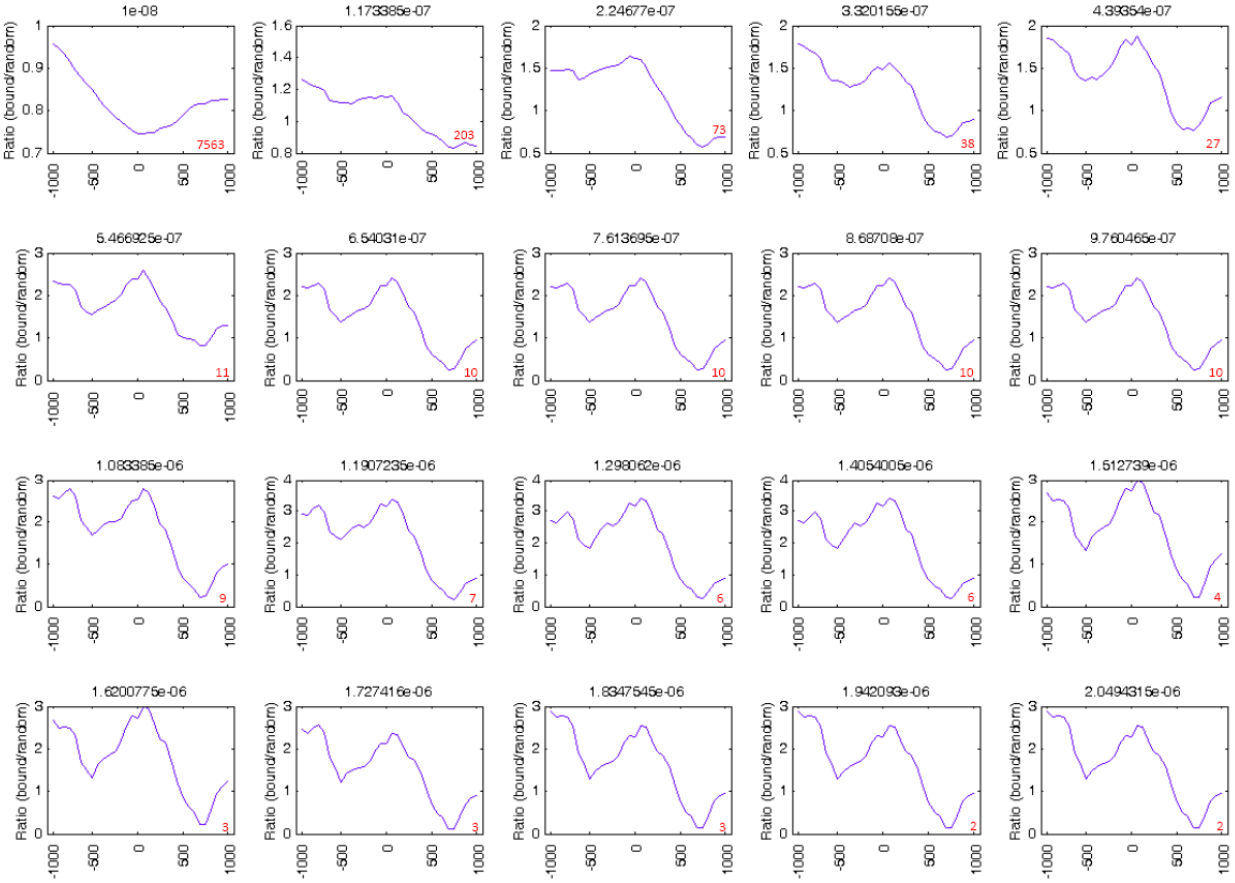
(C) Hnf4a primary motif enrichment within ‘bound’ genomic regions



(D) Hnf4a secondary motif enrichment within ‘bound’ genomic regions



(E) Bcl6b primary motif enrichment within ‘bound’ genomic regions



(F) Bcl6b secondary motif enrichment within ‘bound’ genomic regions