doi:10.1038/nature13994

Table Legends for Supplementary Tables

This file contains description of Supplementary Tables: those that are results-related are in individual files (Supplementary_Table1.xlsx to Supplementary_Table14.xlsx), while those that are methods-related are in a single file Supplementary Tables 15-21.xlsx, with each table in a separate Excel sheet.

Supplementary Table 1 | Number of reads and peak calls for all STARR-seq screens. Total mapped reads, unique fragments (reads that were unique with regard to chromosome, start, end, and strand information and passed the non-heuristic redundancy filter [see ref. 32 and Methods]), and number of peaks for each experiment. The dCP and hsp70 STARR-seq data are from ref. 12 yet have been re-analyzed using a new data processing pipeline. 'Specific peaks' are those that in one cell type are found exclusively with either the dCP or the hkCP but not with both.

Supplementary Table 2 | Gene ontology (GO) analysis in S2 cells, using 'closest TSS' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in S2 cells via 'closest TSS' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 3 | GO analysis in S2 cells, using '1kb TSS' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in S2 cells via '1kb TSS' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 4 | GO analysis in S2 cells, using 'gene loci' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in S2 cells via 'gene loci' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 5 | Gene expression analysis in S2 cells. Gene expression analysis of hkCP and dCP-specific enhancers in S2 cells. Shown are the gene set classifications, as well as a score that combines the over- and under-representation (enrichment/depletion; score = log10 [*P*-value under-representation] - log10 [*P*-value over-representation]).

Supplementary Table 6 | Reproducibility of STARR-seq screens. Pearson correlation coefficients (PCCs) for biological replicates across the entire genome (GW), or between the genome-wide screen and a screen performed for selected genomic regions (~5MB; BAC screens) calculated at peak summit positions. Note that the DSCP and hsp70 screen are from ref. 12.

Supplementary Table 7 | Functional similarity of DSCP and the core promoter of *even-skipped* (*eve*). The DSCP is a 137 nt synthetic core promoter derived from the core promoter of *even-skipped* (*eve*)[11]. We assessed the functional similarity of the DSCP, its 137 nt long wildtype counterpart from the *eve* locus, and a version defined identically to all other core promoter used here (-50 to +50 nt around the TSS) by STARR-seq screens with libraries derived from 29 different BACs containing a total of about 5MB of *Drosophila melanogaster* genomic DNA. All 3 screens were highly similar with Pearson correlation coefficients at peak summit positions above 0.88.

Supplementary Table 8 | GO analysis in OSCs, using 'closest TSS' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in OSCs via 'closest TSS' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value

over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 9 | GO analysis in OSCs, using '1kb TSS' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in OSCs via '1kb TSS' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 10 | GO analysis in OSCs, using 'gene loci' enhancer-to-gene assignment. GO analysis of hkCP and dCP-specific enhancers in OSCs via 'gene loci' enhancer-to-gene assignment strategy. Shown are the GO categories, as well as log10 (*P*-value under-representation) - log10 (*P*-value over-representation). The table is alphabetically sorted according to the GO categories.

Supplementary Table 11 | Motif analysis of specific enhancers in S2 cells. Motif analysis of hkCP and dCP-specific enhancers in S2 cells. Shown are motif names and cluster identification numbers (cluster IDs) that indicate groups of highly similar motifs, log2 value of motif count in enhancer regions over control regions, as well as *P*-values of over- and under-representation. The table is numerically sorted according to the cluster IDs.

Supplementary Table 12 | Motif analysis of specific enhancers in OSCs. Motif analysis of hkCP and dCP-specific enhancers in OSCs. Shown are motif names and cluster identification numbers (cluster IDs) that indicate groups of highly similar motifs, log2 value of motif count in enhancer regions over control regions, as well as *P*-values of over- and under-representation. The table is numerically sorted according to the cluster IDs.

Supplementary Table 13 | Motif analysis of distally-located specific enhancers in S2 cells. Motif analysis of distally-located hkCP and dCP-specific in

S2 cells. Shown are motif names and cluster identification numbers (cluster IDs) that indicate groups of highly similar motifs, log2 value of motif count in enhancer regions over control regions, as well as *P*-values of over- and under-representation. The table is numerically sorted according to the cluster IDs.

Supplementary Table 14 | Motif analysis of distally-located specific enhancers in OSCs. Motif analysis of distally-located hkCP and dCP-specific in OSCs. Shown are motif names and cluster identification numbers (cluster IDs) that indicate groups of highly similar motifs, log2 value of motif count in enhancer regions over control regions, as well as *P*-values of over- and under-representation. The table is numerically sorted according to the cluster IDs.

Supplementary Table 15 | Core promoters used in STARR-seq construct. Core promoter sequences used in the STARR-seq screens and their genomic coordinates, the genes from which they were derived, the biological functions of these genes as well as the core promoter motifs (also called core promoter elements) contained within the core promoters. Core promoter motifs in DSCP and hsp70 are shown as annotated in refs. 11, 45.

Supplementary Table 16 | Bacterial artificial chromosomes (BACs) used for BAC screens. List of all BACs used to clone the BAC STARR-seq library, which together cover about 5MB of the *Drosophila melanogaster* genome. Included are the genomic coordinates of each of the covered regions (assembly release: dm3) and the BAC identifiers.

Supplementary Table 17 | Primers and coordinates of enhancer candidates for luciferase validations. Primer pairs used to amplify enhancer candidates for luciferase validation. Genomic coordinates, the strand orientation of the enhancer candidates as well as the directionality of the contained Transcription Start Site (TSS) are also shown.

Supplementary Table 18 | DRE mutations. DRE mutant sequences used for validation of hkCP dependency on DRE motifs.

Supplementary Table 19 | Enhancer reprogramming. GAGA-to-DRE mutant sequences used to test the sufficiency of DRE to reprogram dCP enhancers, sequence of DRE arrays to test the sufficiency of DRE motifs to activate hkCP core promoter, as well as primer pairs used to amplify dCP enhancers with flanking DRE motifs.

Supplementary Table 20 | Primers and coordinates for 5'RACE. Primers used in 5'RACE of STARR-seq vectors, the enhancers used for each core promoter as well as their strand orientation.

Supplementary Table 21 | Position weight matrices of core promoter element. Log-odds matrices of core promoter elements derived from previously published sources[8,16], together with the *P*-value cut-offs and the regions surrounding the TSSs that were scored.

Supplementary Table 22 | Position weight matrices of TF motifs and core promoter elements from *de novo* motif discovery. Motifs discovered *de novo* in enhancer regions identified using STARR-seq and in core promoter regions around the respective closest annotated transcription starting site (TSS) using MEME[36] (version 4.9.0).