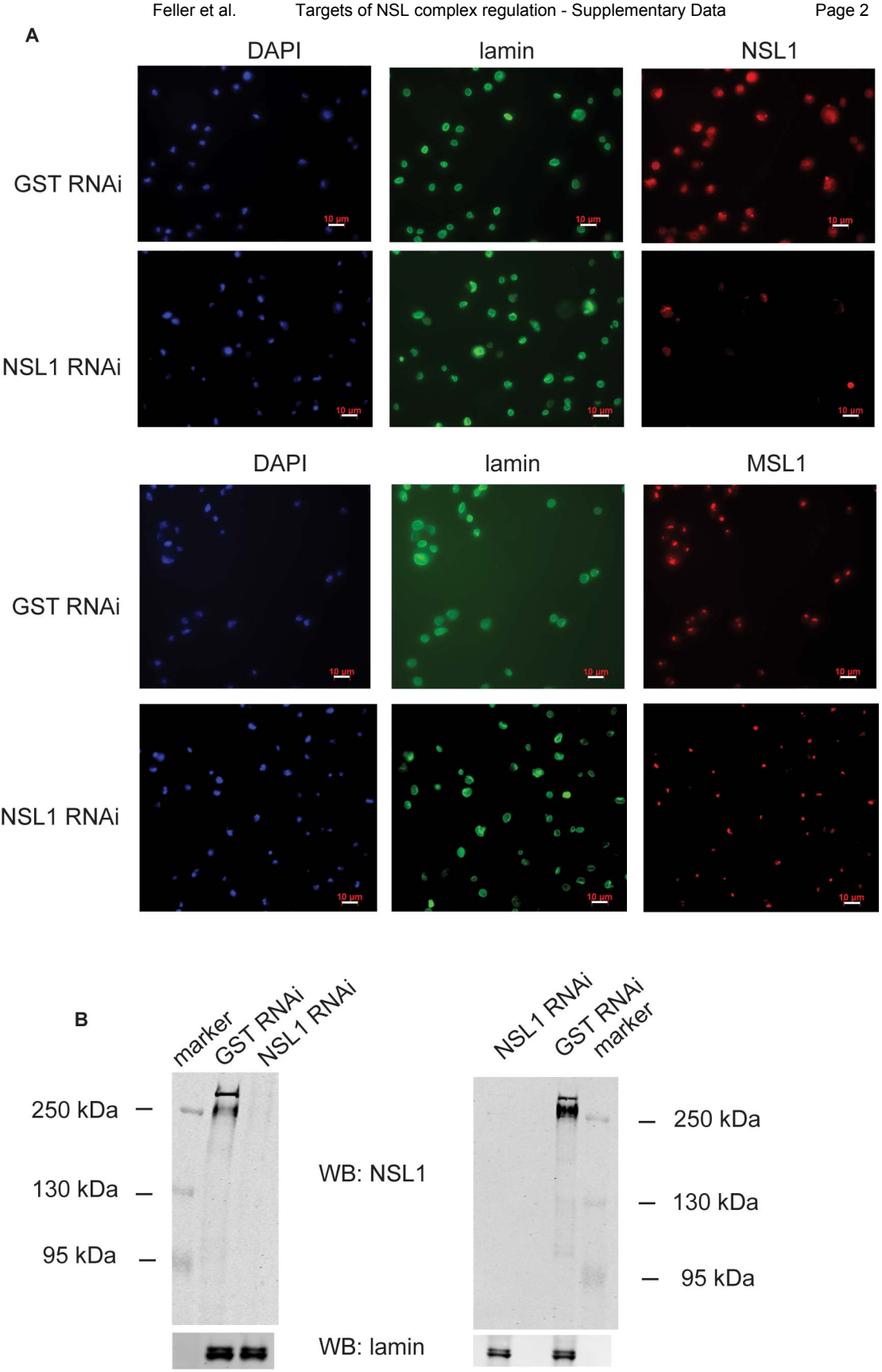
# Supplementary Data

## The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset

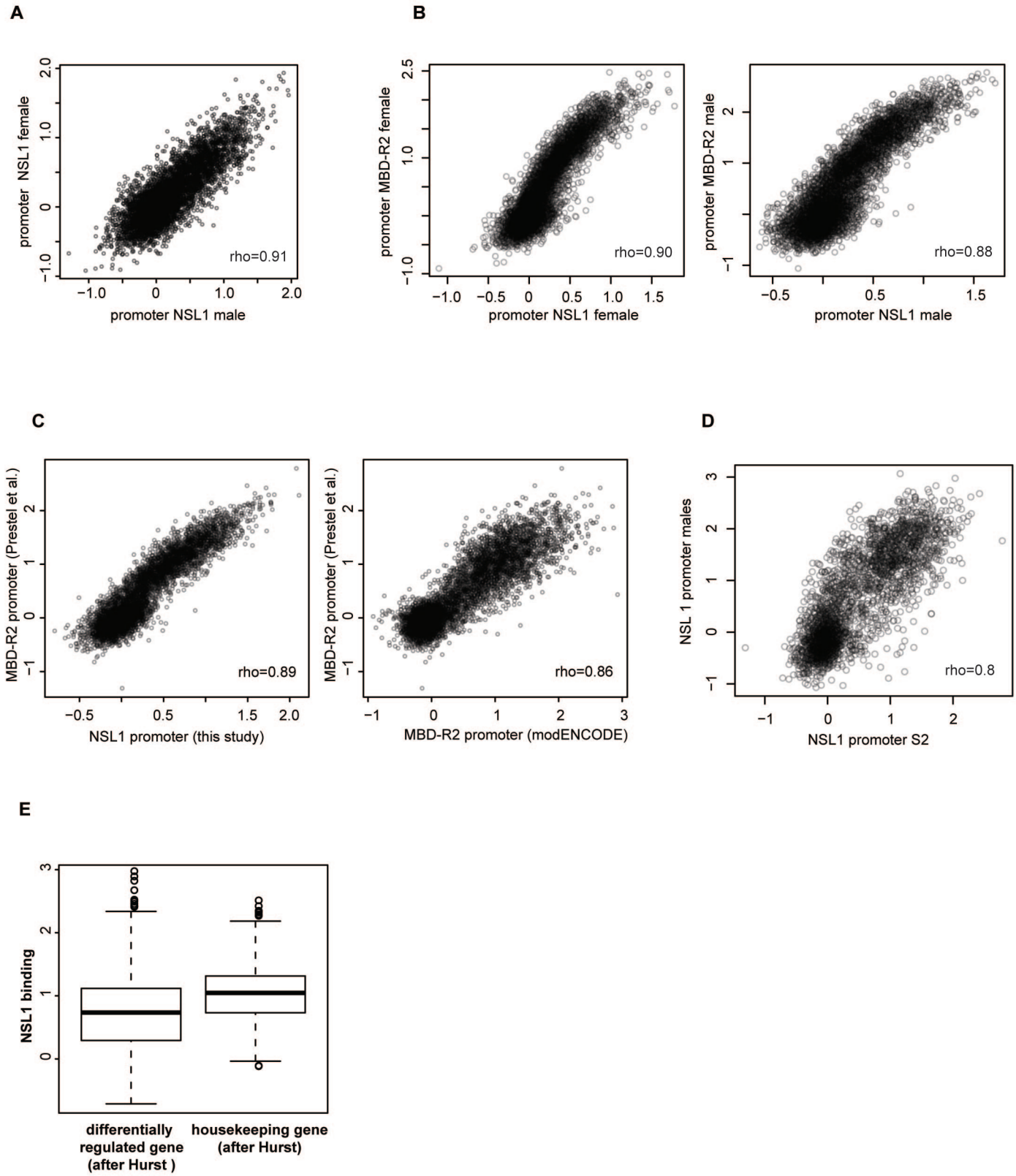
Christian Feller[1], Matthias Prestel[1], Holger Hartmann[2] , Tobias Straub[1], Johannes Söding[2] and Peter B. Becker[1]

[1]Adolf-Butenandt-Institute and [2]Gene Center / Department of Biochemistry, Center for Integrated Protein Science Munich of the Ludwig-Maximilians-University, Munich, Germany

**Corresponding author**
Peter B. Becker
Adolf-Butenandt-Institute,
Schillerstrasse 44, 80336 München, Germany
Phone: 089-2180-75-427
Fax: 089-2180-75-425
Email: pbecker@med.uni-muenchen.de




**Supplementary Figure 1: NSL1 antibody specificity control.**
(A) Indirect immunofluorescence with the indicated antibodies in S2 cells treated with control and NSL1 RNAi. MSL1 (subunit of DCC complex) and lamin stainings are not compromised under NSL1 RNAi conditions. DNA was counterstained with DAPI. Scale bar (10 µm) is indicated in white.
(B) Immunoblot analysis on S2 chromatin extracts (left) and Kc chromatin extracts (right) from NSL1 RNAi, and control RNAi (Glutathion-S-Transferase RNAi) using the NSL1 antibody.
Lamin immunoblot was used as a loading control. The NSL1 protein shows aberrant mobility (11).




**Supplementary Figure 2: NSL complex members bind similar across conditions.**
(A-D) Scatter plots comparing NSL1 promoter binding in adult male and female flies (A), NSL1 vs. MBD-R2 in adult female flies (B, right), NSL1 binding with MBD-R2 binding in adult male flies (B, left), NSL1 binding and MBD-R2 binding (12) in S2 cells (C, left), MBD-R2 binding in S2 cells between the datasets of (12) and the modENCODE consortium (C, right), NSL1 binding in S2 cells with NSL1 binding in adult male cells (D). Rank-based Spearman correlation coefficients (rho) are indicated.
(E) Genes were categorized in “housekeeping” and “differentially regulated” genes, respectively, based on the classification of Hurst and colleagues (22). Binding of NSL1 was displayed in these two groups and was found to be significantly higher in housekeeping genes than differentially regulated genes (Welch two sample t test, $p < 2.2$ e-16). Similar results were obtained using the S2 MBD-R2 ChIP-chip dataset from (12) and the MBD-R2 dataset generated by the modENCODE consortium (17). Note that the “housekeeping” definition after Hurst and colleagues is more stringent, which results in a number of genes classified as “housekeeping” by our definition but “differentially regulated” by Hurst and colleagues.

Supplementary Figure 3: Gene Ontology (GO) enriched biological processes of NSL complex bound genes

| Term | PValue |
|---|---|
| GO:0051188~cofactor biosynthetic process | 2,33E-12 |
| GO:0031023~microtubule organizing center organization | 7,03E-11 |
| GO:0051186~cofactor metabolic process | 1,40E-03 |
| GO:0007017~microtubule-based process | 1,70E-03 |
| GO:0051297~centrosome organization | 1,70E-03 |
| GO:0009108~coenzyme biosynthetic process | 3,00E-03 |
| GO:0007098~centrosome cycle | 4,50E-03 |
| GO:0000226~microtubule cytoskeleton organization | 4,90E-03 |
| GO:0070271~protein complex biogenesis | 5,80E-03 |
| GO:0006461~protein complex assembly | 5,80E-03 |
| GO:0006732~coenzyme metabolic process | 9,70E-03 |
| GO:0016071~mRNA metabolic process | 9,80E-03 |
| GO:0043069~negative regulation of programmed cell death | 9,90E-03 |
| GO:0060548~negative regulation of cell death | 9,90E-03 |
| GO:0006397~mRNA processing | 1,03E-02 |
| GO:0022402~cell cycle process | 1,17E-02 |
| GO:0000278~mitotic cell cycle | 1,17E-02 |
| GO:0051656~establishment of organelle localization | 1,50E-02 |
| GO:0033043~regulation of organelle organization | 1,55E-02 |
| GO:0007010~cytoskeleton organization | 1,82E-02 |
| GO:0006352~transcription initiation | 2,25E-02 |
| GO:0043066~negative regulation of apoptosis | 2,26E-02 |
| GO:0022403~cell cycle phase | 2,40E-02 |
| GO:0007049~cell cycle | 2,46E-02 |
| GO:0051298~centrosome duplication | 2,63E-02 |
| GO:0007173~epidermal growth factor receptor signaling pathway | 2,69E-02 |
| GO:0000279~M phase | 2,76E-02 |
| GO:0051130~positive regulation of cellular component organization | 2,99E-02 |
| GO:0009314~response to radiation | 3,30E-02 |
| GO:0017038~protein import | 3,41E-02 |
| GO:0033365~protein localization in organelle | 3,44E-02 |
| GO:0006367~transcription initiation from RNA polymerase II promoter | 4,07E-02 |
| GO:0006605~protein targeting | 4,08E-02 |
| GO:0051640~organelle localization | 4,21E-02 |
| GO:0006733~oxidoreduction coenzyme metabolic process | 4,92E-02 |
| GO:0048477~oogenesis | 5,56E-02 |
| GO:0016044~membrane organization | 5,59E-02 |
| GO:0006366~transcription from RNA polymerase II promoter | 5,62E-02 |
| GO:0007447~imaginal disc pattern formation | 5,73E-02 |
| GO:0007052~mitotic spindle organization | 5,77E-02 |
| GO:0006468~protein amino acid phosphorylation | 5,88E-02 |
| GO:0031124~mRNA 3'-end processing | 6,00E-02 |
| GO:0007276~gamete generation | 6,06E-02 |
| GO:0019953~sexual reproduction | 6,74E-02 |
| GO:0007051~spindle organization | 6,99E-02 |
| GO:0051276~chromosome organization | 7,19E-02 |
| GO:0008156~negative regulation of DNA replication | 7,27E-02 |
| GO:0007292~female gamete generation | 8,03E-02 |
| GO:0016310~phosphorylation | 8,32E-02 |
| GO:0007349~cellularization | 8,38E-02 |
| GO:0000289~nuclear-transcribed mRNA poly(A) tail shortening | 8,59E-02 |
| GO:0000288~nuclear-transcribed mRNA catabolic process, deadenylation-d | 8,59E-02 |
| GO:0035222~wing disc pattern formation | 8,77E-02 |
| GO:0051493~regulation of cytoskeleton organization | 8,77E-02 |
| GO:0009628~response to abiotic stimulus | 9,11E-02 |
| GO:0032504~multicellular organism reproduction | 9,29E-02 |
| GO:0048609~reproductive process in a multicellular organism | 9,29E-02 |
| GO:0051301~cell division | 9,52E-02 |
| GO:0006281~DNA repair | 9,65E-02 |
| GO:0065003~macromolecular complex assembly | 9,66E-02 |
| GO:0006259~DNA metabolic process | 9,70E-02 |
| GO:0016072~rRNA metabolic process | 9,70E-02 |
| GO:0006612~protein targeting to membrane | 9,70E-02 |
| GO:0006364~rRNA processing | 9,70E-02 |

Gene Ontology (GO) enrichment analysis was performed for biological processes using DAVID GO SLIM (57).
P values were derived from a Fisher Exact test using the DAVID web-implementation.
All NSL1 bound genes were compared to all active genes in S2 cells.

Supplementary Figure 3

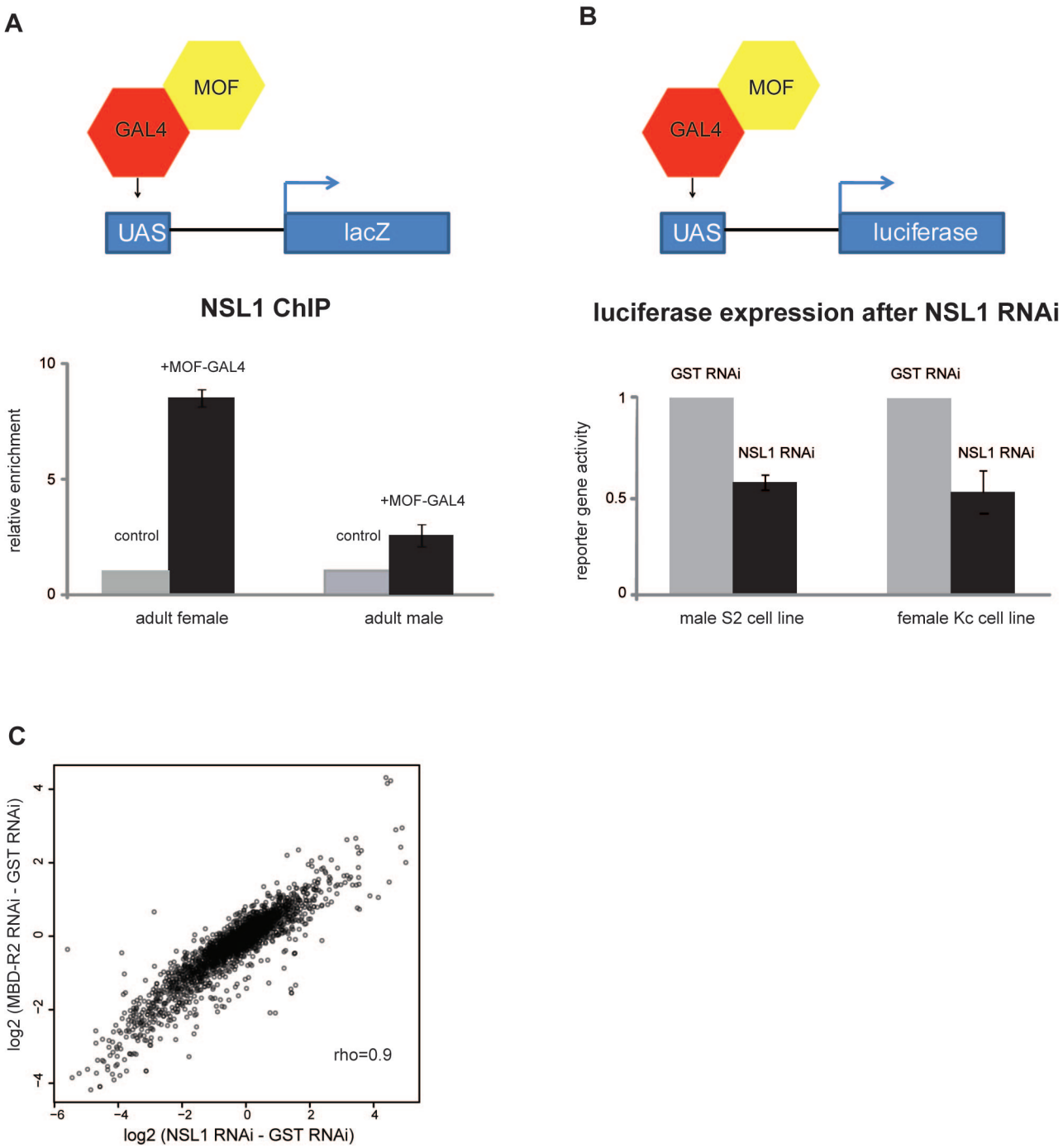
# Supplementary Figure 4: Gene Ontology analysis of active NSL1 unbound genes



| Term | PValue |
|---|---|
| GO:0007606~sensory perception of chemical stimulus | 1,57E-07 |
| GO:0055114~oxidation reduction | 4,48E-07 |
| GO:0007600~sensory perception | 3,79E-06 |
| GO:0050890~cognition | 6,59E-06 |
| GO:0007155~cell adhesion | 1,49E-05 |
| GO:0022610~biological adhesion | 1,49E-05 |
| GO:0035218~leg disc development | 1,79E-05 |
| GO:0006952~defense response | 2,93E-05 |
| GO:0048569~post-embryonic organ development | 5,89E-05 |
| GO:0035286~leg segmentation | 1,03E-04 |
| GO:0035285~appendage segmentation | 1,03E-04 |
| GO:0046942~carboxylic acid transport | 1,03E-04 |
| GO:0015849~organic acid transport | 1,03E-04 |
| GO:0006865~amino acid transport | 1,35E-04 |
| GO:0015837~amine transport | 1,35E-04 |
| GO:0007166~cell surface receptor linked signal transduction | 1,41E-04 |
| GO:0007552~metamorphosis | 1,44E-04 |
| GO:0035111~leg joint morphogenesis | 1,92E-04 |
| GO:0007444~imaginal disc development | 2,05E-04 |
| GO:0045087~innate immune response | 3,64E-04 |
| GO:0007186~G-protein coupled receptor protein signaling pathway | 3,97E-04 |
| GO:0048707~instar larval or pupal morphogenesis | 3,99E-04 |
| GO:0048563~post-embryonic organ morphogenesis | 4,28E-04 |
| GO:0007560~imaginal disc morphogenesis | 4,28E-04 |
| GO:0042742~defense response to bacterium | 5,24E-04 |
| GO:0048737~imaginal disc-derived appendage development | 5,54E-04 |
| GO:0048732~gland development | 5,60E-04 |
| GO:0009886~post-embryonic morphogenesis | 6,30E-04 |
| GO:0048736~appendage development | 6,39E-04 |
| GO:0006955~immune response | 6,55E-04 |
| GO:0035214~eye-antennal disc development | 7,53E-04 |
| GO:0035114~imaginal disc-derived appendage morphogenesis | 8,79E-04 |
| GO:0007447~imaginal disc pattern formation | 9,21E-04 |
| GO:0048565~gut development | 9,88E-04 |
| GO:0035107~appendage morphogenesis | 0,00101 |
| GO:0016052~carbohydrate catabolic process | 0,00139 |
| GO:0009617~response to bacterium | 0,00175 |
| GO:0009791~post-embryonic development | 0,00192 |
| GO:0050830~defense response to Gram-positive bacterium | 0,00215 |
| GO:0007354~zygotic determination of anterior/posterior axis, embryo | 0,00215 |
| GO:0035108~limb morphogenesis | 0,00221 |
| GO:0060173~limb development | 0,00221 |
| GO:0035110~leg morphogenesis | 0,00221 |
| GO:0007157~heterophilic cell adhesion | 0,00238 |
| GO:0007449~proximal/distal pattern formation, imaginal disc | 0,00238 |
| GO:0035220~wing disc development | 0,00258 |
| GO:0007431~salivary gland development | 0,00261 |
| GO:0035272~exocrine system development | 0,00261 |
| GO:0035120~post-embryonic appendage morphogenesis | 0,00276 |
| GO:0042067~establishment of ommatidial polarity | 0,00283 |
| GO:0010623~developmental programmed cell death | 0,00378 |
| GO:0007450~dorsal/ventral pattern formation, imaginal disc | 0,00433 |
| GO:0007610~behavior | 0,00438 |
| GO:0035215~genital disc development | 0,00444 |
| GO:0007219~Notch signaling pathway | 0,00492 |
| GO:0007366~periodic partitioning by pair rule gene | 0,00509 |
| GO:0001708~cell fate specification | 0,00608 |
| GO:0019731~antibacterial humoral response | 0,00617 |
| GO:0050877~neurological system process | 0,00683 |
| GO:0048190~wing disc dorsal/ventral pattern formation | 0,00808 |
| GO:0007362~terminal region determination | 0,00834 |
| GO:0050909~sensory perception of taste | 0,00834 |
| GO:0009954~proximal/distal pattern formation | 0,00834 |
| GO:0006006~glucose metabolic process | 0,00866 |
| GO:0006928~cell motion | 0,00924 |
| GO:0016337~cell-cell adhesion | 0,00933 |
| GO:0002165~instar larval or pupal development | 0,00964 |
| GO:0009065~glutamine family amino acid catabolic process | 0,01053 |
| GO:0006584~catecholamine metabolic process | 0,01053 |
| GO:0009636~response to toxin | 0,01053 |
| GO:0040034~regulation of development, heterochronic | 0,01053 |
| GO:0018958~phenol metabolic process | 0,01053 |
| GO:0034311~diol metabolic process | 0,01053 |
| GO:0009712~catechol metabolic process | 0,01053 |

| Term | PValue |
|---|---|
| GO:0007451~dorsal/ventral lineage restriction, imaginal disc | 0,01073 |
| GO:0048598~embryonic morphogenesis | 0,01169 |
| GO:0044275~cellular carbohydrate catabolic process | 0,01187 |
| GO:0046164~alcohol catabolic process | 0,01187 |
| GO:0006508~proteolysis | 0,01269 |
| GO:0007476~imaginal disc-derived wing morphogenesis | 0,01297 |
| GO:0007494~midgut development | 0,01349 |
| GO:0009996~negative regulation of cell fate specification | 0,01349 |
| GO:0042659~regulation of cell fate specification | 0,01349 |
| GO:0010453~regulation of cell fate commitment | 0,01349 |
| GO:0010454~negative regulation of cell fate commitment | 0,01349 |
| GO:0007626~locomotory behavior | 0,01432 |
| GO:0007472~wing disc morphogenesis | 0,01436 |
| GO:0043473~pigmentation | 0,01489 |
| GO:0007164~establishment of tissue polarity | 0,0158 |
| GO:0001736~establishment of planar polarity | 0,0158 |
| GO:0007435~salivary gland morphogenesis | 0,01581 |
| GO:0022612~gland morphogenesis | 0,01581 |
| GO:0001654~eye development | 0,01737 |
| GO:0019318~hexose metabolic process | 0,01837 |
| GO:0009072~aromatic amino acid family metabolic process | 0,01867 |
| GO:0035167~larval lymph gland hemopoiesis | 0,01867 |
| GO:0006935~chemotaxis | 0,01867 |
| GO:0007469~antennal development | 0,01867 |
| GO:0007440~foregut morphogenesis | 0,01867 |
| GO:0035166~post-embryonic hemopoiesis | 0,01867 |
| GO:0007608~sensory perception of smell | 0,01867 |
| GO:0001738~morphogenesis of a polarized epithelium | 0,02213 |
| GO:0006959~humoral immune response | 0,02258 |
| GO:0008219~cell death | 0,02319 |
| GO:0016265~death | 0,02319 |
| GO:0007411~axon guidance | 0,02345 |
| GO:0035289~posterior head segmentation | 0,02396 |
| GO:0042688~crystal cell differentiation | 0,02396 |
| GO:0008587~imaginal disc-derived wing margin morphogenesis | 0,02585 |
| GO:0043067~regulation of programmed cell death | 0,02652 |
| GO:0010941~regulation of cell death | 0,02652 |
| GO:0007478~leg disc morphogenesis | 0,02674 |
| GO:0019730~antimicrobial humoral response | 0,02721 |
| GO:0042060~wound healing | 0,02983 |
| GO:0012501~programmed cell death | 0,03034 |
| GO:0009310~amine catabolic process | 0,03038 |
| GO:0006007~glucose catabolic process | 0,03224 |
| GO:0019320~hexose catabolic process | 0,03224 |
| GO:0048066~pigmentation during development | 0,03293 |
| GO:0016054~organic acid catabolic process | 0,03293 |
| GO:0046395~carboxylic acid catabolic process | 0,03293 |
| GO:0006576~biogenic amine metabolic process | 0,03535 |
| GO:0050829~defense response to Gram-negative bacterium | 0,03535 |
| GO:0006334~nucleosome assembly | 0,03535 |
| GO:0035222~wing disc pattern formation | 0,03789 |
| GO:0030030~cell projection organization | 0,03796 |
| GO:0000902~cell morphogenesis | 0,03819 |
| GO:0005976~polysaccharide metabolic process | 0,03902 |
| GO:0046365~monosaccharide catabolic process | 0,03963 |
| GO:0042386~hemocyte differentiation | 0,04036 |
| GO:0035162~embryonic hemopoiesis | 0,04036 |
| GO:0000272~polysaccharide catabolic process | 0,04036 |
| GO:0048568~embryonic organ development | 0,04036 |
| GO:0048749~compound eye development | 0,04235 |
| GO:0046667~compound eye retinal cell programmed cell death | 0,04283 |
| GO:0035161~imaginal disc lineage restriction | 0,04283 |
| GO:0009253~peptidoglycan catabolic process | 0,04283 |
| GO:0006027~glycosaminoglycan catabolic process | 0,04283 |
| GO:0000270~peptidoglycan metabolic process | 0,04283 |
| GO:0006206~pyrimidine base metabolic process | 0,04283 |
| GO:0035287~head segmentation | 0,04283 |
| GO:0035223~leg disc pattern formation | 0,04283 |
| GO:0007479~leg disc proximal/distal pattern formation | 0,04283 |
| GO:0014070~response to organic cyclic substance | 0,04416 |
| GO:0007419~ventral cord development | 0,04416 |
| GO:0043279~response to alkaloid | 0,04416 |
| GO:0006096~glycolysis | 0,04557 |
| GO:0045596~negative regulation of cell differentiation | 0,0459 |
| GO:0042981~regulation of apoptosis | 0,04595 |
| GO:0043068~positive regulation of programmed cell death | 0,04734 |
| GO:0030031~cell projection assembly | 0,04734 |
| GO:0010942~positive regulation of cell death | 0,04734 |

Gene Ontology (GO) enrichment analysis was performed using active but NSL1 unbound genes against the background of all active genes. DAVID GO SLIM (57) terms were used for biological processes. P values were derived from a Fisher Exact test using the DAVID web-implementation.

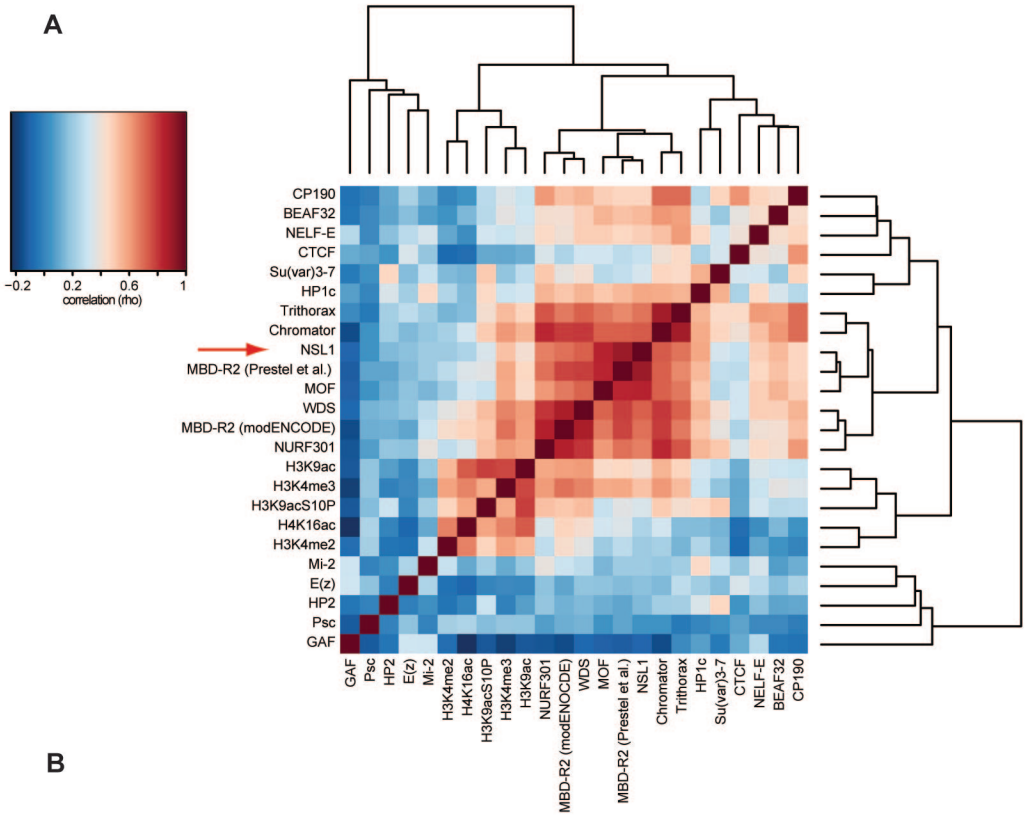


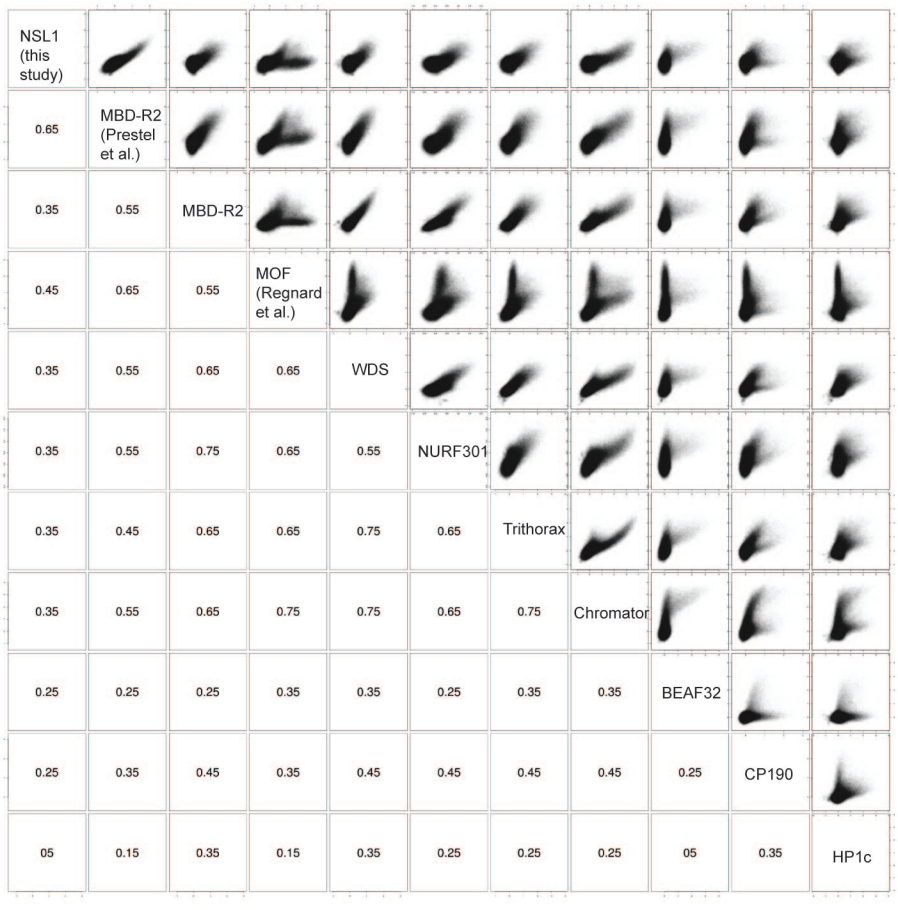

**Supplementary Figure 5: NSL1 stimulates transcription on a reporter gene and genome-wide.**
(A) NSL1 binding correlates with stronger activation in female flies as compared to male flies after MOF recruitment to a reporter gene. We previously showed that recruitment of the MOF-GAL4 fusion protein to an UAS-element upstream of the lacZ reporter gene leads to stronger activation in adult female flies as compared to adult male flies (12). MBD-R2 recruitment as measured by ChIP at the lacZ gene follows this trend (12). Here, we tested NSL1 binding at the reporter gene and find a similar recruitment of NSL1 which is significantly ($p < 0.05$, 2-sided unpaired t-test) increased in adult female flies when compared to adult male flies. We conclude that the higher activation potential of MOF in female flies correlates with an increased NSL complex recruitment. Error bars represent SEM of at least three independent biological replicates. (B) RNAi mediated ablation of NSL1 leads to a diminished MOF-dependent activation of a luciferase reporter gene in male S2 cells and female Kc cells. Transiently co-transfection of MOF-GAL4 fusion gene, UAS-luciferase reporter and renilla reporter (used for normalization) lead to an activation of the luciferase reporter. NSL1 RNAi in male S2 and female Kc cells significantly diminished ($p < 0.05$, 2-sided unpaired t-test) the luciferase reporter expression when compared to a control RNAi treatment. Error bars represent SEM of at least three independent biological replicates. (C) Transcriptome changes are similar in response to NSL1 RNAi and MBD-R2 RNAi. Scatter plot of transcriptome changes after NSL1 (this study) and MBD-R2 RNAi (12), respectively, relative to control RNAi. Rank-based Spearman correlation coefficient (rho) is indicated in the corner.

Supplementary Figure 6: Gene Ontology enrichment analysis of NSL1 RNAi downregulated genes

| Term | Pvalue |
|---|---|
| GO:0006259~DNA metabolic process | 1,30E-05 |
| GO:0006281~DNA repair | 2,70E-04 |
| GO:0034660~ncRNA metabolic process | 9,00E-04 |
| GO:0006412~translation | 7,23E-02 |
| GO:0033554~cellular response to stress | 0,001 |
| GO:0006399~tRNA metabolic process | 0,001 |
| GO:0006974~response to DNA damage stimulus | 0,001 |
| GO:0008033~tRNA processing | 0,002 |
| GO:0034470~ncRNA processing | 0,005 |
| GO:0006284~base-excision repair | 0,006 |
| GO:0006366~transcription from RNA polymerase II promoter | 0,011 |
| GO:0006367~transcription initiation from RNA polymerase II promoter | 0,012 |
| GO:0006351~transcription, DNA-dependent | 0,014 |
| GO:0006352~transcription initiation | 0,015 |
| GO:0032774~RNA biosynthetic process | 0,016 |
| GO:0006396~RNA processing | 0,017 |
| GO:0009451~RNA modification | 0,023 |
| GO:0006260~DNA replication | 0,032 |

Downregulated genes (fold change < -2) were compared to all active genes (Affymetrix log2-scale > 4).
Gene Ontology enrichment analysis was performed using DAVID (57), testing for biological processes using the GO SLIM terms. P values were derived from a Fisher Exact test using the DAVID web-implementation.
A similar result was obtained using the R/Bioconductor package GOstats (data not shown).



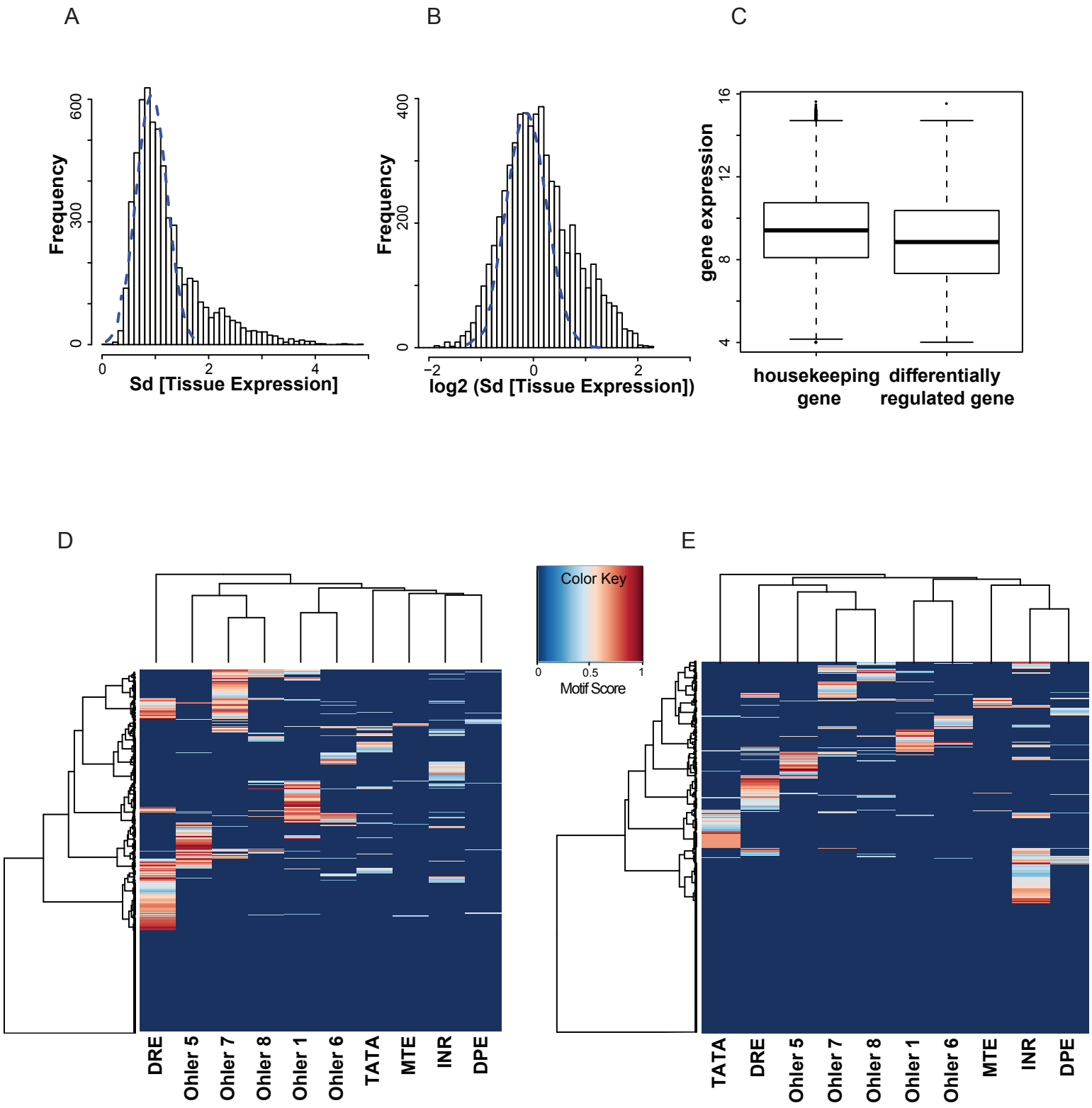




**Supplementary Figure 7: Co-localization of NSL complex members with Trithorax and NURF301.**
(A) Heatmap was generated as in Figure 3A but included in addition to the profiles shown in Figure 3A the ChIP-chip datasets of MBD-R2 (12), MBD-R2 (modENCODE) and MOF (modENCODE).
(B) Probe-wise scatter plot indicated co-localization of the NSL complex with Trithorax and NURF301. ChIP-chip datasets were used from the modENCODE consortium database (17) unless otherwise indicated. Probe-wise enrichments were calculated, smoothed and genomic positions were compared in 100 base pair intervals across samples. Note that all probes are shown including the inactive unbound probes and the X chromosomal active probes where NSL1 and MBD-R2 binding deviate from MOF binding towards the transcribed regions of genes (compare to (12)).
Spearman correlation coefficient is indicated.

Supplementary Figure 8: Summary of factor and sequence occupancies across different conditions.

| | % Binding to all Genes | Active Genes | Inactive Genes | % Binding to all active Housekeeping | active DRG | Enrichment Active/ Inactive | Enrichment Housekeeping/ /DRG |
|---|---|---|---|---|---|---|---|
| BEAF32 | 17,9 | 33,5 | 1,1 | 39,9 | 17 | 31,7 | 2,3 |
| Ohler 1 | 7,6 | 12,7 | 2,1 | 15 | 6,7 | 5,9 | 2,2 |
| HP2 | 2 | 3,2 | 0,7 | 3,8 | 1,7 | 4,8 | 2,2 |
| DRE | 14,4 | 21,2 | 7,2 | 24,6 | 12,3 | 2,9 | 2 |
| Ohler 5 | 8,6 | 12,1 | 4,7 | 14 | 7,2 | 2,6 | 2 |
| Chromator | 44,7 | 80,7 | 6 | 92,7 | 49,4 | 13,4 | 1,9 |
| Ohler 7 | 10,4 | 15,3 | 5,1 | 17,6 | 9,5 | 3 | 1,9 |
| MBD-R2 | 38,9 | 71,7 | 3,7 | 82,1 | 44,7 | 19,4 | 1,8 |
| WDS | 39,6 | 73,5 | 3,1 | 83,9 | 46,7 | 23,5 | 1,8 |
| Trithorax | 40,6 | 74,4 | 4,4 | 84,7 | 47,5 | 16,7 | 1,8 |
| CP190 | 29 | 47,5 | 9,2 | 54 | 30,5 | 5,2 | 1,8 |
| NURF301 | 42,2 | 77,5 | 4,2 | 87,8 | 51,1 | 18,4 | 1,7 |
| Ohler 6 | 6 | 7,6 | 4,2 | 8,5 | 5,1 | 1,8 | 1,7 |
| HP1c | 15,5 | 29,1 | 1 | 32,6 | 20 | 29,1 | 1,6 |
| H3K4me3 | 47,5 | 87,1 | 5,1 | 94,9 | 67 | 17,2 | 1,4 |
| Su(var)3-7 | 1,8 | 2,6 | 1,1 | 2,8 | 2 | 2,4 | 1,4 |
| H3K9ac | 46,9 | 84,5 | 6,6 | 92 | 65,3 | 12,8 | 1,4 |
| Nelf-E | 28,8 | 54,9 | 0,9 | 59,3 | 43,3 | 59,6 | 1,4 |
| H4K16ac | 50,0 | 80,0 | 17,5 | 78,9 | 58,6 | 9,7 | 1,3 |
| Nelf-B | 30,7 | 58,6 | 0,9 | 62,7 | 47,7 | 62,3 | 1,3 |
| CTCF | 4 | 4,4 | 3,5 | 4,5 | 4,2 | 1,2 | 1,1 |
| Ohler 8 | 5,8 | 6,8 | 4,7 | 6,9 | 6,6 | 1,4 | 1 |
| TATA | 9,7 | 8,2 | 11,4 | 7,9 | 8,9 | 0,7 | 0,9 |
| E(z) | 1,3 | 0,8 | 1,9 | 0,7 | 1,1 | 0,4 | 0,6 |
| INR | 18,4 | 15,6 | 21,5 | 12,7 | 23,2 | 0,7 | 0,5 |
| H3.3 | 20,8 | 32,3 | 8,4 | 26,2 | 48,3 | 3,8 | 0,5 |
| MTE | 2 | 1,9 | 2,2 | 1,5 | 3 | 0,9 | 0,5 |
| Psc | 1,6 | 1,1 | 2 | 0,8 | 1,9 | 0,6 | 0,4 |
| RING | 2,9 | 3,3 | 2,4 | 2,4 | 5,8 | 1,4 | 0,4 |
| DPE | 3,6 | 3,4 | 3,8 | 2,5 | 6 | 0,9 | 0,4 |
| Mi-2 | 2,6 | 4,7 | 0,3 | 2,9 | 9,2 | 15,8 | 0,3 |
| GAF | 8,3 | 14,8 | 1,4 | 8,6 | 30,7 | 10,6 | 0,3 |
| H3K27me3 | 35,8 | 4,1 | 69,8 | 1,9 | 9,8 | 0,1 | 0,2 |
| Pc | 2 | 0,5 | 3,6 | 0,2 | 1,2 | 0,1 | 0,1 |

Factor and sequence occupancy is given relative to all genes in the *D. melanogaster* genome (FlyBase dmel5.32 annotation) (column 2), all active genes (column 3, gene activity measured by modENCODE polymerase II dataset), all inactive genes (column 4), all active housekeeping genes (column 5) and all active “differentially regulated” genes (DRG, column 6). Column 7 indicates the ratio of active genes over inactive genes, and column 8 indicates the ratio of active housekeeping genes over active “differentially regulated” genes. The list was sorted according to the enrichment towards active housekeeping genes (column 8) and overrepresented and underrepresented factors are labeled in red and green, respectively. Only modENCODE ChIP-chip profiles were used.

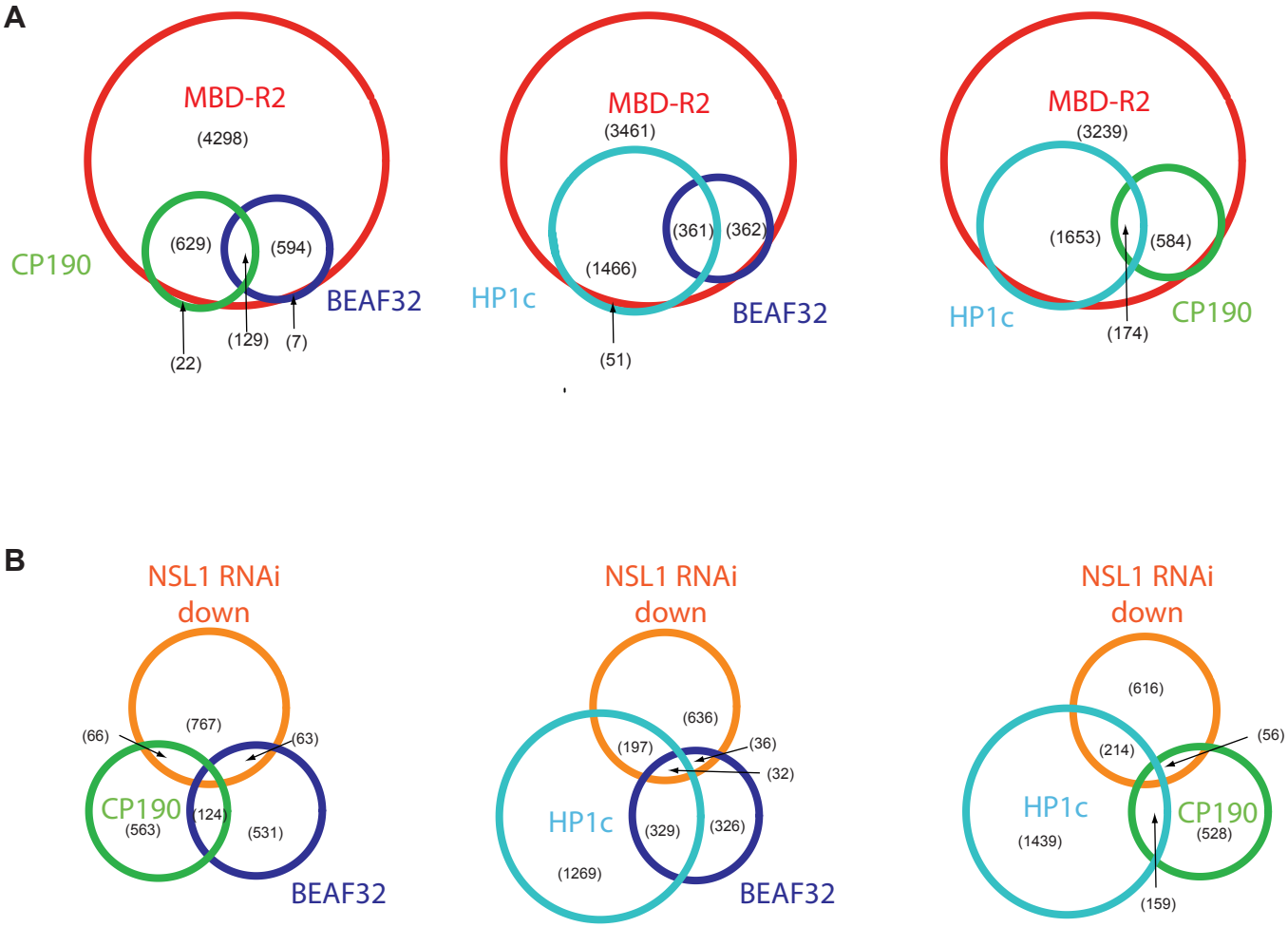



**Supplementary Figure 9: Characterization of core promoter motifs at active genes.**
(A) Histogram depicting the distribution of standard deviations of active genes across 40 *Drosophila* tissues (21). Normal distribution is overlayed with a blue dashed line. (B) Same as (A) but log2 transformed standard deviations. (C) Comparison of gene expression values in housekeeping genes (left box) and differentially regulated genes (right box). Gene expression values were taking from the log2 transformed Affymetrix control RNAi transcriptome dataset. (D) Core promoter motif distribution of active housekeeping genes. Heatmap presentation of motif score (columns) along active housekeeping genes (rows). The dendrogram to the right was generated by hierarchical clustering of gene promoters according to their motif score using the 'complete' method implemented in the R package 'hclust'. The dendrogram above the heatmap was generated by hierarchical clustering ('ward method') of the Spearman correlated motif scores. Note that 1269 of the 4523 active housekeeping genes do not show any significant signature of at least one of the investigated core promoter motifs. See Material and Methods for detailed information on the definition of the 'motif score'. (E) Same as (D) but for active differentially regulated genes. Note that 480 of the 1367 active differentially regulated genes do not show any significant signature of at least one of the investigated core promoter motifs.

**Supplementary Figure 10: NSL1 peaks are centered over core promoter motifs characteristic for housekeeping genes**

| motif | total number sequences enriched for motif | sequence fraction enriched for motif |
|---|---|---|
| DRE | 1106 | 19.9 |
| Ohler 1 | 703 | 12.7 |
| Ohler 5 | 556 | 10.0 |
| Ohler 7 | 510 | 9.2 |
| Ohler 6 | 251 | 4.5 |
| Ohler 8 | 109 | 2.0 |
| INR | 84 | 1.5 |
| TATA | 38 | 0.7 |
| MTE | 33 | 0.6 |
| DPE | 12 | 0.2 |

De-novo sequence search (Hartmann and Soeding, in preparation) of the sequence surrounding +/- 50 base pairs centered on the NSL1 ChIP-Seq (11) peak summit. Column 2 indicates the number of identified sequences containing the respective motif. Column 3 displays the fractions of each identified motif relative to all identified motifs.




**Supplementary Figure 11: HP1c and insulator binding proteins co-localize with the NSL complex at non-functional sites.** (A) Venn diagram representation of gene sets bound by the NSL complex (using the MBD-R2 ChIP-chip profile generated by the modENCODE consortium) and heterochromatin protein 1c (HP1c), CP190 and BEAF32. (B) Venn diagram representation of gene sets downregulated after NSL1 RNAi and bound by the heterochromatin protein 1c (HP1c), CP190 and BEAF32. Numbers in brackets represent the number of genes in the respective group (for clearity numbers with little information content are not shown). We used the modENCODE MBD-R2 data set as a proxy for the NSL complex to minimize technical variation in the comparison with the BEAF32/CP190/HP1c datasets. Similar results were obtained using the NSL1 dataset and the MBD-R2 dataset (12), respectively.