Supplementary Data for:

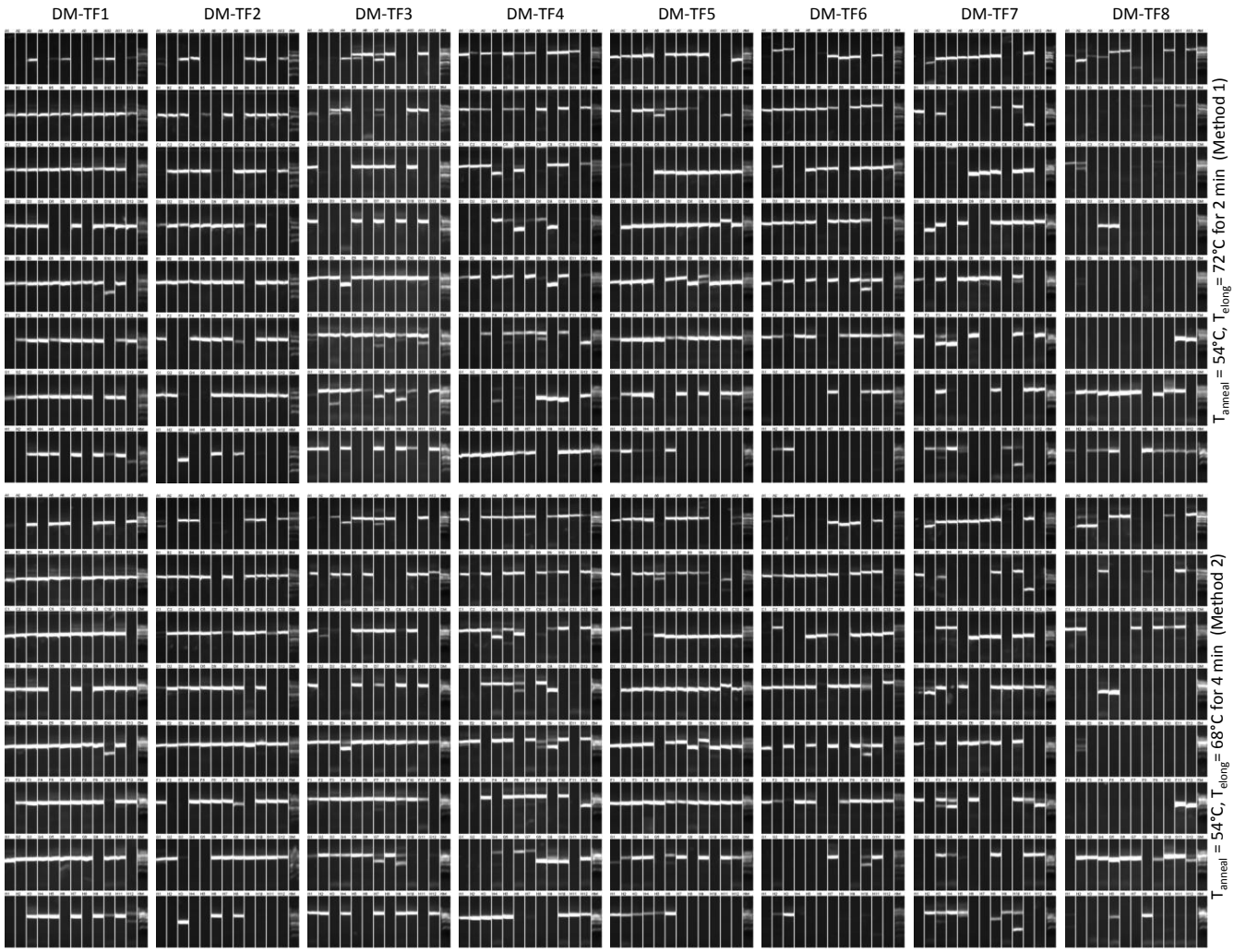
iSLIM: a comprehensive approach to mapping and characterizing gene regulatory networks

Sylvie Rockel[1], Marcel Geertz[2], Korneel Hens[3], Bart Deplancke[3] & Sebastian J. Maerkl[1,*]

[1] Laboratory of Biological Network Characterization, Institute of Bioengineering, School of Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2] Department of Molecular Biology, University of Geneva, Geneva, Switzerland
[3] Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

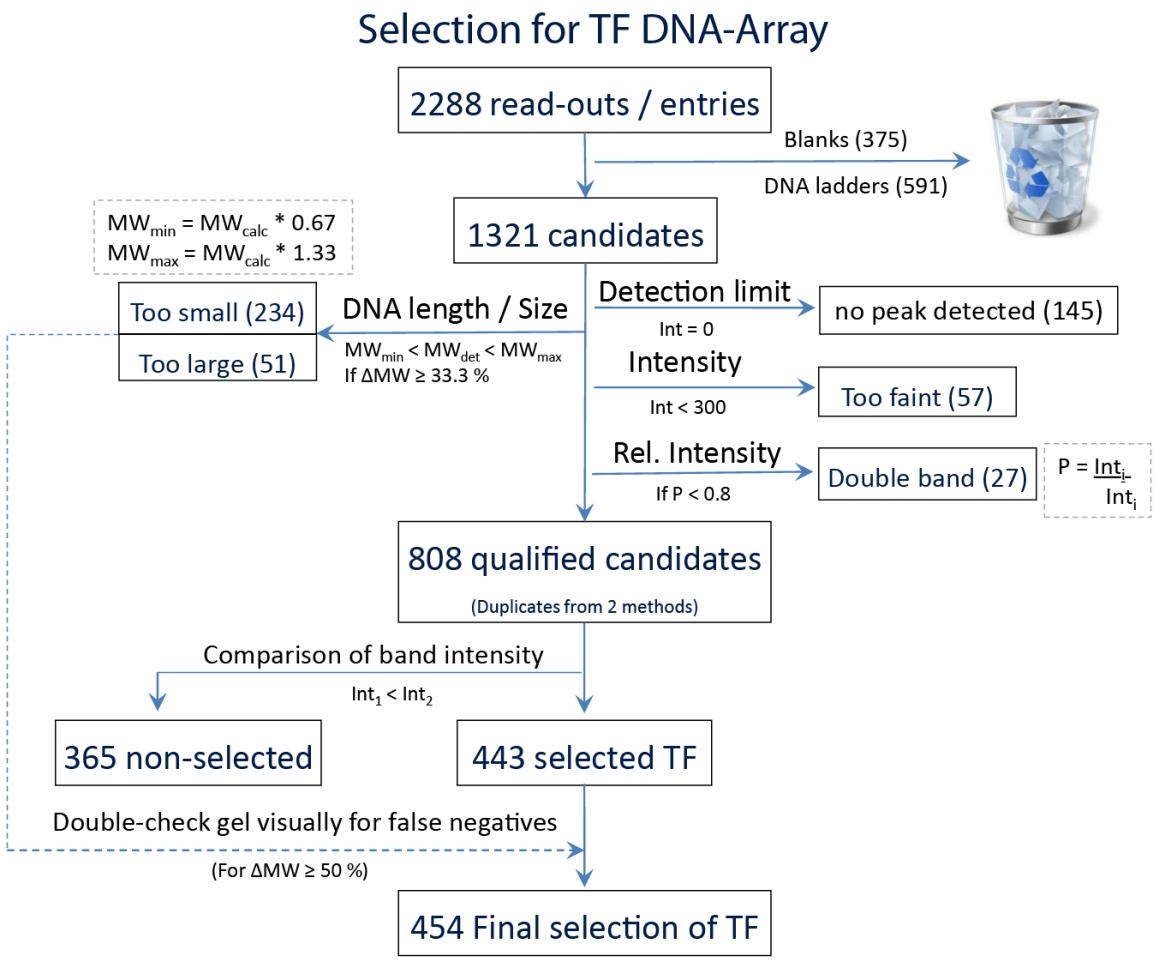
* To whom correspondence should be addressed. Tel: +41 21 693 7835; Fax: +41 21 693 7830; Email: sebastian.maerkl@epfl.ch

Supplementary Figure S1: 96-well agarose gels.



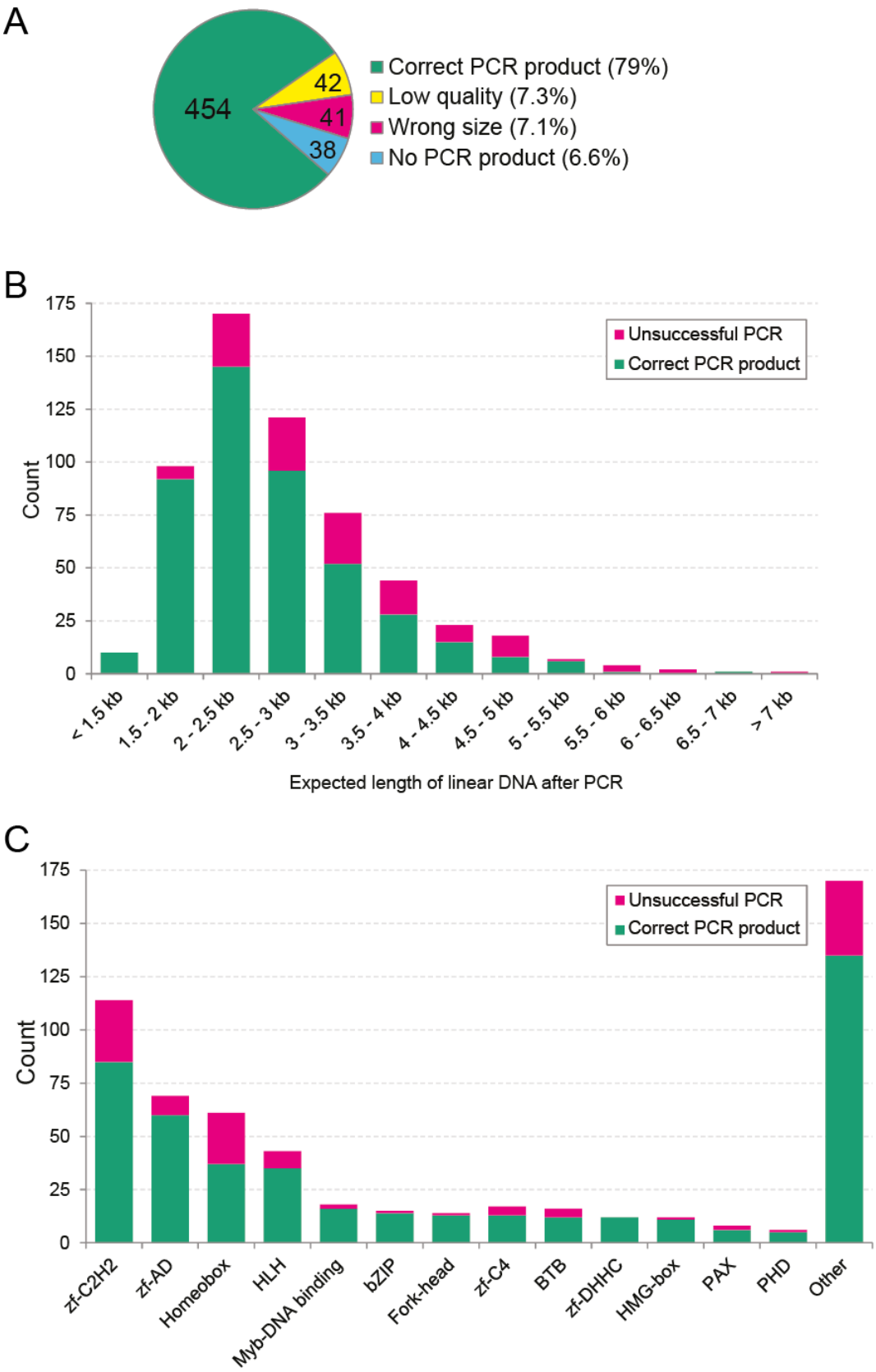
Two PCR reactions per sample were automatically assessed for quality with a MatLab-based image analysis program, which returned the number of detected DNA fragments for each gel lane, its size and intensity.

Supplementary Figure S2: Flow chart of selection process of PCR products.



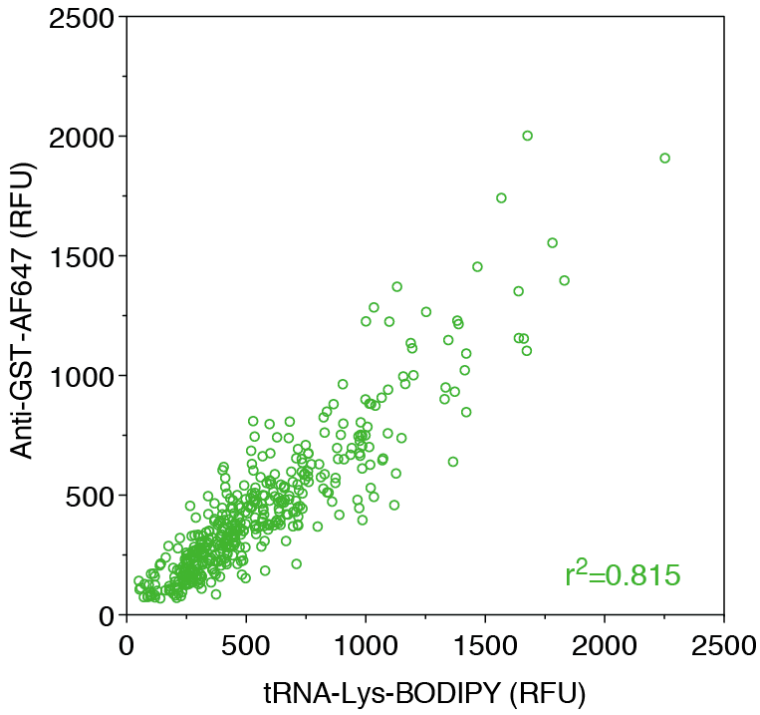
PCR read-outs, which did not match our selection criteria for size (± 33 % from expected size), band intensity and purity of the PCR product, were excluded from the candidate pool. For samples with successful PCR products in both methods we selected the sample with the higher intensity value on the electrophoresis gel. We double-checked for false negatives within the pool of size-excluded candidates for samples without a positive read-out from both methods.

Supplementary Figure S3: PCR results.



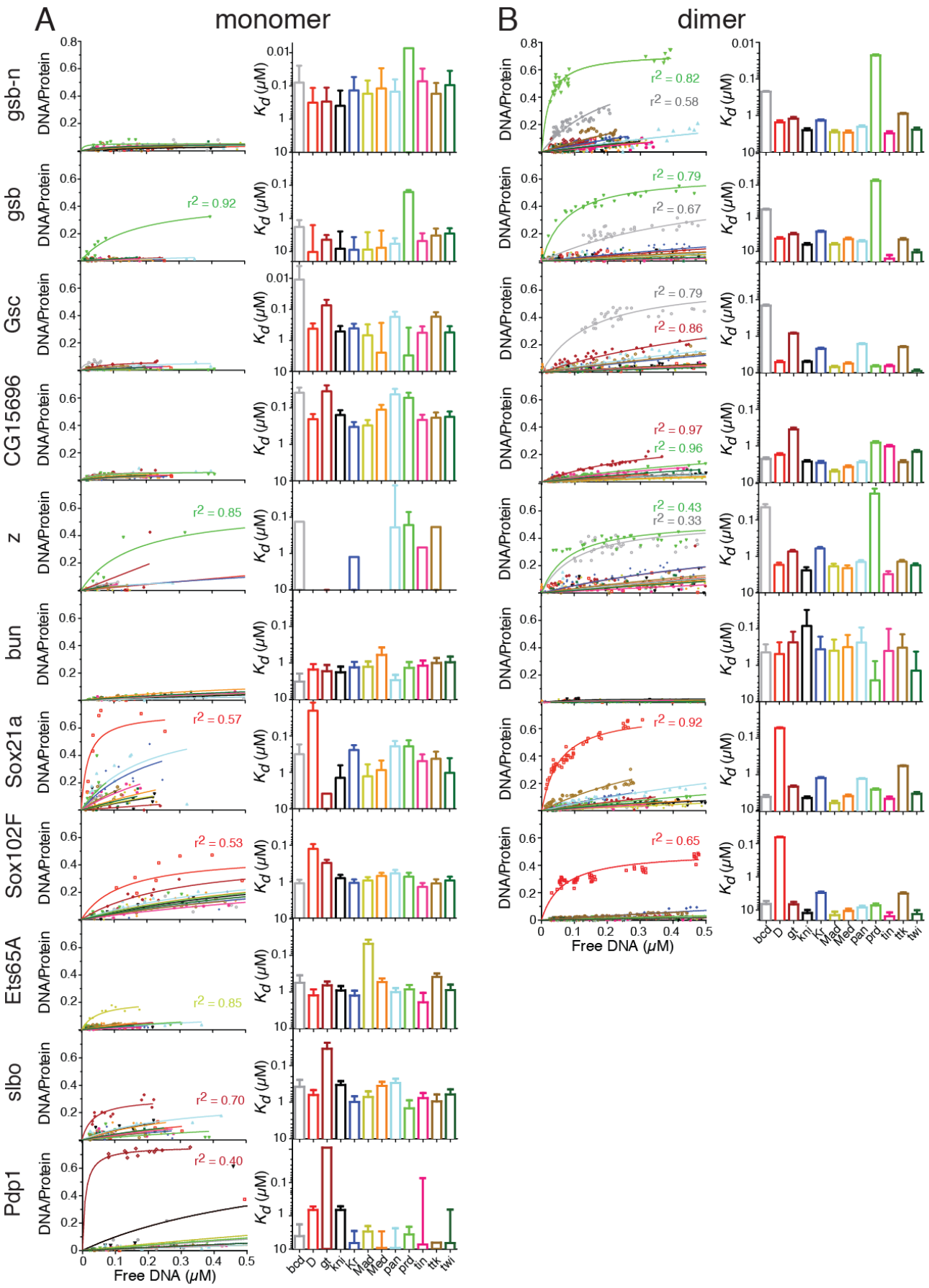
(A) Pie chart showing the PCR performance after quality assessment. (b,c) histograms of PCR success rate of (B) expected DNA fragment size, and (C) DNA-binding protein domains within the Drosophila TFs.

Supplementary Figure S4: Correlation between detected signals for protein expression.



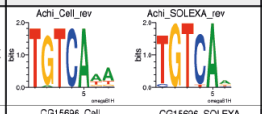
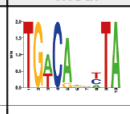
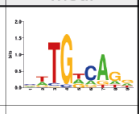
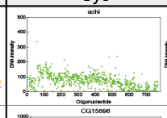
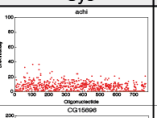
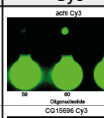
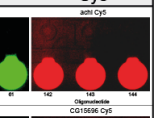
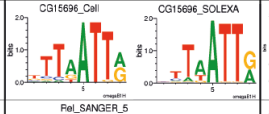
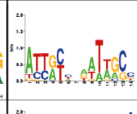
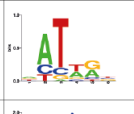
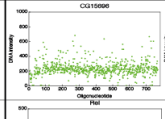
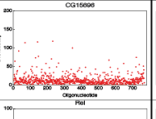
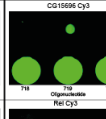
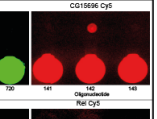
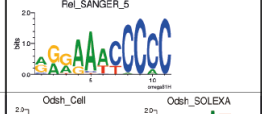
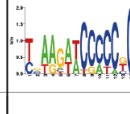
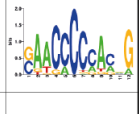
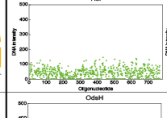
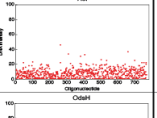
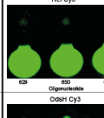
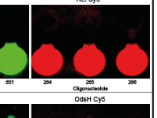
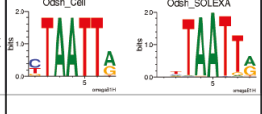
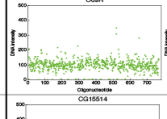
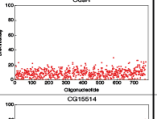
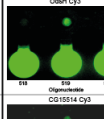
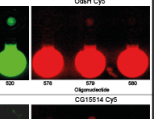
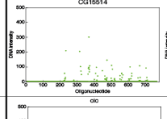
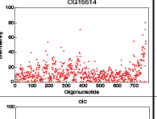
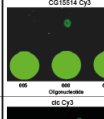
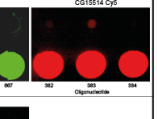
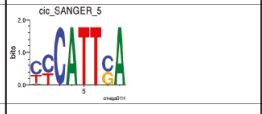
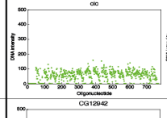
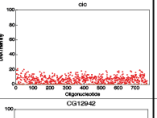
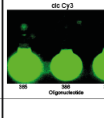
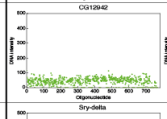
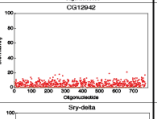
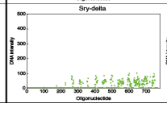
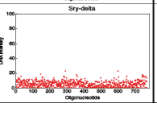
Protein expression levels were characterized by relating BODIPY intensities to the number of surface-bound TFs carrying a lysine-BODIPY-charged tRNA. Correlation between signal intensities of protein pull-down labeled with a lysine-BODIPY-charged tRNA and a secondary GST-antibody (Hilyte Fluor™ 647) on the same TF array.

Supplementary Figure S5: Specificity and affinity measurements of TFs to DNA sequences.



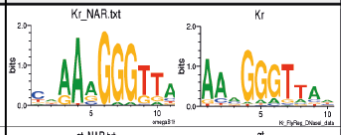
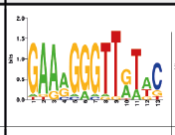
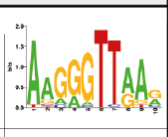
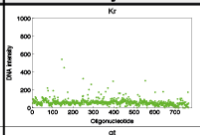
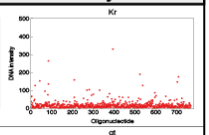
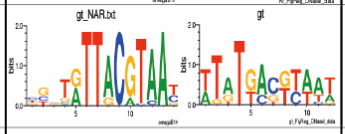
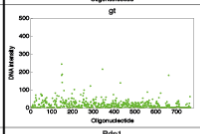
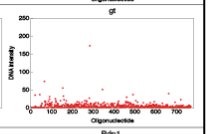
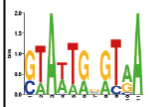
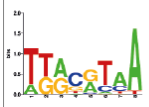
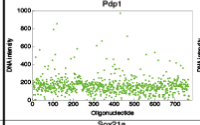
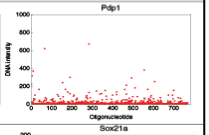
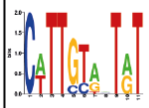
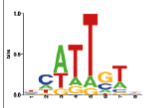
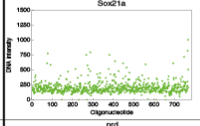
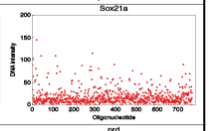
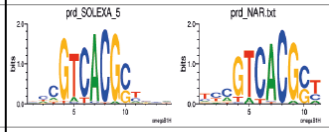
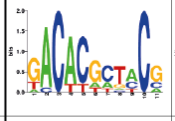
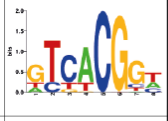
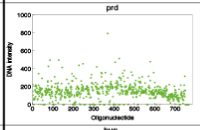
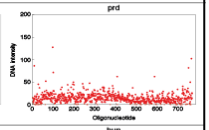
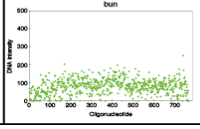
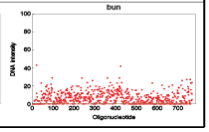
Specificity and affinity measurements of TFs to DNA sequences for all 11 newly identified TFs, which bound to one or more of the 12 DNA consensus motifs (bcd, D, gt, kni, Kr, Mad, Med, pan, prd, tin, ttk, twi). As in Fig. 5 concentration-dependent binding curves are derived from fitting the measured DNA/Protein ratio signals over available consensus DNA (in solution) to a single-site binding model (left); binding affinities to each consensus motif are plotted as Kd values with standard errors (right).
DNA binding was measured in two different modes: For monomer-DNA binding, target DNA could bind only to surface-immobilized TFs. In the second mode, TFs were allowed to interact with DNA consensus sequences in solution prior to antibody binding and detection, which potentially allowed for TF binding to DNA as dimers.

Supplementary Figure S6: Analysis of 8 randomly chosen TFs.

TF (MW)	DBD family	Previous motifs	MEME motif	iSLIM motif	Intensity plots Cy3	Intensity plots Cy5	Array images Cy3	Array images Cy5
achi (75.6 kDa)	Homeobox							
CG15696 (48.2 kDa)	Homeobox							
Rel (137.7 kDa)	RHD							
OdsH (70.6 kDa)	Homeobox		-	-				
CG15514 (63.3 kDa)	zf-BED	-	-	-				
cic (178.0 kDa)	HMG-box		-	-				
CG12942 (106.8 kDa)	zf-C_2H_2	-	-	-				
Sry-delta (77.9 kDa)	zf-C_2H_2	-	-	-				

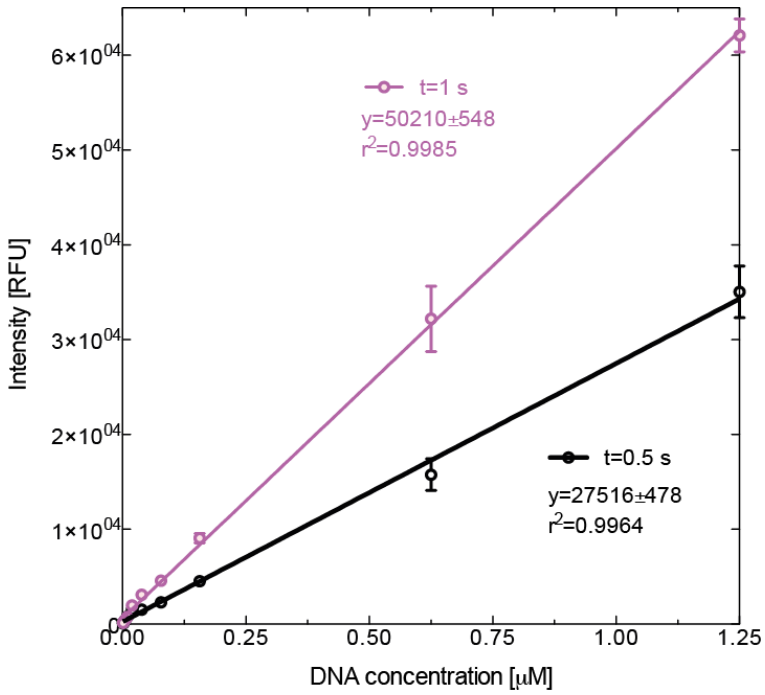
Binding ability of randomly chosen TFs was assessed on an oligonucleotide array containing all possible 65,532 8mer sequences (de Bruijn library). The table summarizes plots of measured Cy3 and Cy5 intensities after background correction for each oligonucleotide and a supporting scan image for each channel for detected TF-DNA binding. Motifs found with MEME & iSLIM, as well as previously known motifs (if any) are shown as sequence logos.

Supplementary Figure S7: Analysis of the selected TFs.

TF (MW)	DBD family	Previous motifs	MEME motif	iSLIM motif	Intensity plots Cy3	Intensity plots Cy5
Kr (82.7 kDa)	zf-C_2H_2					
gt (77.1 kDa)	bZIP_2		-	-		
Pdp1 (58.3 kDa)	bZIP_1	-				
Sox21a (68.7 kDa)	HMG-box	-				
prd (90.8 kDa)	Paired homeo-box					
bun (47.6 kDa)	Leucine zipper	-				

Binding ability of selected TFs was assessed on an oligonucleotide array containing all possible 65,532 8mer sequences (de Bruijn library). Table summarizes plots of measured Cy3 and Cy5 intensities after background correction for each oligonucleotide for detected TF-DNA binding. Motifs found with MEME & iSLIM, as well as previously known motifs (if any) are shown as sequence logos.

Supplementary Figure S8: Cy5 DNA Calibration curve



Measured Cy5 DNA chamber intensities were fitted as a function of known concentration of a serial dilution of soluble DNA (mean ± standard deviation) to a linear regression model for two exposure times.

Supplementary Table T1: Primer sequences of target DNA with motif sequence highlighted.

TF	Primer sequence
bcd	5'-CGC**GGATTAG**CTCCGGCGGTATGAC-3'
D	5'-CGC**GTCCATTGTTCT**CTCCGGCGGTATGAC-3
gt	5'-CGC**ATTACGTAATAA**CTCCGGCGGTATGAC-3'
kni	5'-CGC**AAAACTAGAGCAA**CTCCGGCGGTATGAC-3'
Kr	5'-CGC**TAACCCTTTTG**CTCCGGCGGTATGAC-3'
Mad	5'-CGC**GCTGCCGGCGCGGC**CTCCGGCGGTATGAC-3'
Med	5'-CGC**AACAGGCGAAA**CTCCGGCGGTATGAC-3'
pan	5'-CGC**CTTTGATC**CTCCGGCGGTATGAC-3'
prd	5'-CGC**CCAATTTGTCACGCT**CTCCGGCGGTATGAC-3'
tin	5'-CGC**CTCAAGTG**CTCCGGCGGTATGAC-3'
ttk	5'-CGC**ATTATCCTGG**CTCCGGCGGTATGAC-3'
twi	5'-CGC**CCCGCATATGTTG**CTCCGGCGGTATGAC-3'

Supplementary Table T2: Summary of recent gene-centric methods.

Assay	Species	Clone Coverage	Proteins Expressed	Quality Control	Reference
iSLIM	*D. melanogaster*	60% (454 of 755 predicted TFs)	93% (423 of 454) (full-length)	37.5% (3/8) TFs returned PWMs 62.5% (5/8) TFs bound DNA 44% (4/9) literature curated interactions recovered	this work
Protein Array	*H. sapiens*	80% (1370 TF clones) (4191 clones in total)	90% (180 of 200) (full-length)	15% (201/1370) TFs returned PWMs 41.2% - 34.5% TFs bound DNA	Hu et al., Cell 2009
Y1H	*D. melanogaster*	78% (588 of 755 predicted TFs)	not known	26% (5/19) literature curated interactions recovered	Hens et al., Nature Methods 2011
Y1H	*H. sapiens*	69% (988 of 1,434 predicted TFs)	not known	eY1H: 16% (5/31) literature curated interactions recovered Y1H diploid: 35% (11/31) Y1H haploid: 65% (20/31) Combined: 77% (24/31)	Reece-Hoyes et al., Nature Methods 2011
eY1H	*C. elegans*	89% (834 of 937 TFs in compendium)	not known	55% (5/9) ChIP interactions recovered	Reece-Hoyes et al., Nature Methods 2011
eY1H	*A. thaliana*	654 clones 92% of TFs expressed in the root stele 74.5% of TFs expressed in the root	not known	50% (7/14) previously confirmed interactions recovered (physical or regulatory in *planta*)	Gaudinier et al., Nature Methods 2011