Supplementary Note A. - Comparison of algorithmic design choices

Computational identification of ChIP enriched regions corresponding to protein-DNA interactions from mapped reads is a crucial step in the analysis of ChIP-Seq data. Similarly to Qeseq, the other 14 algorithms we compare are generally based on three main modules: (i) construction of signal profiles, (ii) localization of events and (iii) filtering of candidate results.

I. Signal processing strategies

All ChIP-Seq algorithms construct an initial representation of the distribution of sequence reads. The most common method is to use a sliding kernel to estimate the local density, either for both DNA strands together, or for each DNA strand separately (QuEST). The types of kernels in ChIP-Seq analysis are the isotropic box-kernel (CCAT, ChIPDiff, MACS, RSEG, SISSRs and W-ChIPeaks), the anisotropic box-kernel oriented in the read DNA strand direction (ERANGE, PeakSeq, SICER, SWEMBL, and TPIC) and the Gaussian kernel (FSeq, Qeseq and QuEST). The most recent algorithms utilize additional strategies to improve enrichment estimation, for instance Hidden Markov Models (ChIPDiff and RSEG), topological analysis based on shape (TPIC), or iterative re-calibration of signals (Qeseq).

II. Localization of events

Following the estimation of the read distribution, an enrichment indicator is determined in the form of threshold scores to identify putative enrichment events. These scores are based on the value of the read density of either the ChIP signal alone (FSeq, FindPeaks) or a combination of ChIP and control signal densities (CCAT, ChIPDiff, ERANGE, PeakSeq, and QuEST). Alternatively, threshold is implemented by parametric approaches based on Poisson distributions (SISSRs, SICER), local Poisson distributions (MACS), or non-parametric approaches such as the Cramér-von Mises test (Qeseq) and percentile rank statistic (W-ChIPeaks) and tree-shape statistics (TPIC).

In order to identify the genomic boundaries of candidate events, algorithms rely on internal threshold criteria to merge ChIP enriched regions into broader putative events. The merging criterion generally requires stringent adjacency. To address the problems of imperfect mapping and low-coverage, some algorithms apply improvements such as the inclusion of non-enriched stretches (SICER, Qeseq). RSEG employs a genome segmentation model, where the boundaries of events are identified using an HMM model based on the Baum-Welch algorithm.

III. Filtering of candidate events

All algorithms employ ad hoc filters to eliminate false positive events. Available filters fall into three categories: filters of deviation from expectation, multiple testing p-value corrections, and local estimation of specificity. Common filters in the first category assess the imbalance in the number of reads between the DNA strands within the boundaries of an event (ERANGE, Qeseq, QuEST, SICER) and whether the positional distribution of reads corresponding to the negative and positive strands within each event is bimodal (CCAT, ChIPDiff, MACS, QuEST and SISSRs). The other two filter categories depend on event detection in the control signal. Multiple testing corrections are based on implementation of an empirical false discovery rate (FDR) control defined as the ratio between the number of events identified in the ChIP signal and the number of events identified in the control signal (MACS, SISSRs, W-ChIPeaks). Local

assessment of signal specificity verifies that positive events are not overlapping with non-specific signals (Qeseq).

Supplementary Note B. – qPCR validations

I. Is qPCR a gold standard for ChIP-Seq validation?

Comparative studies of ChIP-Seq algorithms rely on datasets derived from quantitative polymerase chain reaction (qPCR) experiments and examine which algorithms predict the validation sites correctly. Although qPCR experiments are relatively inexpensive and can be executed in large batches (~100 sites) in a short amount of time (~50 sites/day in triplicates), there are important differences between ChIP-Seq experiments and ChIP-qPCR based validations.

First, conventional ChIP-qPCR does not require a population amplification step (LM-PCR), which is an essential part of the ChIP-Seq protocol. LM-PCR amplification of complex DNA populations is known for amplification bias, which can give rise to false enrichment in the ChIP-Seq signal. Second, conventional ChIP-qPCR examines a specific position in both the input and the specific ChIP at the same time and therefore accurately measures and compensates for variability in the original chromatin used for the experiment. Third, for economic reasons, most ChIP-Seq experiments do not include IgG control samples. In contrast, conventional ChIP-qPCR experiments generally include an IgG control sample, which gives a reliable measurement of the specificity of the experiment.

Is qPCR based validation reliable and is it independent of the ChIP-Seq signal?

In the design of our comparative study, we contemplated whether qPCR was a suitable gold standard for ChIP-Seq validation, because qPCR does not provide a direct confirmation of the interaction between DNA and the protein of interest. Rather, it is a repetition of the sequencing experiment, often at a coarser resolution. Many artifacts and false interactions stemming from sequencing experiments may therefore also occur again in the validation data, giving a false impression of the biological accuracy of the result.

Comparison between qPCR-based enrichment evaluation and density estimation generated by ChIP-Seq algorithms is acceptable from a signal analysis point of view, as both investigate the same quantity (enrichment of the IP fraction over a control). However, it is not a meaningful validation for the purpose of assessing the biology of the system. In order to obtain an optimal confirmation of the biological question that motivates the use of a ChIP-Seq experiment, ChIP-qPCR should be combined with knockdown or knockout approaches, targeting the transcription factor or enzyme responsible for the histone modification investigated. To our knowledge, no large ChIP-Seq validation dataset is based on these techniques.

The goal of our study was to assess algorithmic performance from a signal analysis perspective. Thus, we surmised that qPCR validations, albeit their experimental limitations, are a valid and acceptable tool.

II. Can we leverage on existing qPCR dataset?

Before performing novel validation experiments we studied the possibility of leveraging on existing datasets where ChIP-Seq experiments were validated by qPCR experiments. Some qPCR-based datasets had too few validated sites to produce any meaningful statistics. Other datasets were designed based on the predictions of a single algorithm, and thus captured only certain types of binding patterns. In addition, the strongest peaks supported by the largest number

of reads are typically chosen for validations, making it difficult to assess in an unbiased fashion the performance of ChIP-Seq algorithms across the full range of enrichment amplitudes. More generally, handpicked validation sites are often biased toward specific biological questions that an investigator is trying to address.

We tested these hypotheses in a preliminary assessment of algorithmic performance using transcription factor (TF) datasets (Supplementary Table 5) for E2F4 in mouse muscle cells (MYO.E2F4.MT, (34)), for STAT1 in HeLa S3 cells (STAT1, (16)) and for SIN3A and SIN3B in mouse muscle cells (MYO.SIN3A.GM, MYO.SIN3B.GM, MYO.SIN3A.MT and MYO.SIN3B.MT, (34)). We derived Equation 4 to compute balanced accuracy (or AUCROC for binary classifiers) based on the law of total probabilities and the results obtained are significantly more consistent and reproducible (Figure 2) compared to using Equation 3. We applied our formula to existing TF datasets, and verified the performance of existing algorithms. Unexpectedly, all algorithms showed poor performances; the AUCROC performance measure for most algorithms was around 0.5 (Supplementary Table 6). However, closer inspection of the validated sites revealed that although the algorithms were properly designed, the poor performance was a reflection of the biased selection of limited sites for qPCR. In fact, these sites were not selected with the purpose of collecting data for algorithmic evaluation and performance. This shortcoming justified the construction of a fair, large and well-designed novel validation dataset. To avoid an overlap with existing efforts (4,5) we designed proper validation experiments for histone modifications.

III. How to design a ChIP-Seq oriented qPCR validation dataset?

We applied our novel VDA tool for selecting sites that maximally discriminate between available algorithms. Several technical issues had to be addressed before we could effectively compare our qPCR validations with the ChIP-Seq data.

For our study, we designed 3 batches of validation experiments such that each validation batch had around 100 high-quality qPCR validations. Since the expected failure rate of qPCR (i.e. when it is not possible to determine whether a site is a true positive or a true negative) is around 15-20%, we actually tested 115 sites for each batch. When identifying loci for validation of ChIP-Seq data derived from MNAse digested cross-linked chromatin, it is important to consider the issue of resolution. Careful titration experiments in our lab have revealed that the maximum amount of cross-linked mono-nucleosomes that can be obtained without over-digesting the DNA is ~85% of the total nucleosome population (Asp, Dynlacht, unpublished observations). The remaining material remains as di-, or tri-nucleosome fragments. Extending the time of MNAse digestion in order to generate a pure mono-nucleosomal population results in significant over-digestion of mono-nucleosomal DNA to a smaller size than the average 150bp of a single nucleosome and is therefore not recommended. The presence of di- and tri- nucleosomal DNA is not an issue for the ChIP-Seq procedure itself since there is a mono-nucleosome size selection step during the library preparation. Therefore, the resulting sequencing data is of mono-nucleosomal resolution. However, conventional ChIP-qPCR experiments using MNAse digested chromatin do not contain a size selection step. Subsequently, the enriched material will consist of a mix of mono- (~150bp), di- (~400bp) or tri-nucleosomal (~750bp) DNA fragments and therefore does not give mono-nucleosomal resolution. A minimal distance of 2-3 nucleosomes

between loci to be validated by ChIP-qPCR is therefore required in order to reliably distinguish between events despite that fact that the ChIP-Seq data may show discrete and separate events.

Supplementary Note C. – VDA protocol for designing ChIP-Seq validation datasets

I. General considerations

Applying the VDA procedure to ChIP-Seq data consists of three steps: 1) defining candidate sites, 2) applying the VDA algorithm to the list of candidate sites, and 3) validating the sites suggested by the VDA algorithm. We designed two procedures to select candidate sites from the compendium of all predicted (positive and negative) sites returned by multiple algorithms.

Following experimental considerations (see Supplementary Notes A.III), we added a size-selection filtering of candidate validation sites between step 1 and step 2. In particular, we selected validation sites in the range of 500-1000bp. The lower limit of 500bp was chosen, as it is larger than the size of DNA fragments obtained by MNase digestion and the upper limit of 1000bp was chosen to guarantee that a single primer could provide sufficient coverage for the interval. These sizes are dataset specific and should be adapted to the length-scales of the system of interest.

The VDA approach was then applied in step 2 to the list of candidate validation sites to generate a list of diverse representative validation sites. To perform the validation experiments (step 3), primers were designed starting from the center of the site. Primer design might not be feasible in certain genomic regions due to thermodynamic barriers preventing primer hybridization, repetitive sequences, sub-telomeric or centromeric regions. When the predicted binding events fell in those genomic locations, we substituted them by other candidate sites that had the same fingerprints (see next section on Genome segmentation protocol).

II. Genome segmentation protocol

We generated the initial list of candidate validation sites by using the events predicted by 15 algorithms that we applied to ChIP-Seq datasets. Specifically, we partitioned the genome into non- overlapping intervals determined by the collection of all the event boundaries predicted by these algorithms (Supplementary Figure 12). For any interval *i* between any pair of adjacent boundaries, a fingerprint was defined as a binary vector $U_i = [A_{1i}, A_{2i}, ..., A_{ni}]$ where A_{ji} was 1 if the interval overlaps with a site predicted by algorithm *j* or 0 otherwise.

To illustrate the concept of a fingerprint, we present a toy example in which there are 5 algorithms (A1, A2, A3, A4 and A5) and an artificial genome consisting of 10 nucleotides. For each nucleotide, the prediction is either 1 (where the given algorithm identifies an event) or 0 (where the given algorithm does not identify an event). The table below represents the predictions of these algorithms for the 10 nucleotides:

Nucleotide	*A1*	*A2*	*A3*	*A4*	*A5*
1	1	1	0	0	0
2	0	0	0	0	0
3	1	1	1	1	1
4	1	0	1	0	0
5	0	1	1	1	1
6	1	1	0	0	0
7	0	1	1	1	1
8	1	0	0	1	1
9	1	1	0	0	0
10	1	0	0	1	1

Based on this table, the fingerprint for nucleotide 1 is the vector v1=[1 1 0 0 0], and for nucleotide 2 is the vector v2=[0 0 0 0 0] and so on.

We then group the nucleotides that have identical fingerprints:

Nucleotide	*A1*	*A2*	*A3*	*A4*	*A5*
1	**1**	**1**	**0**	**0**	**0**
6	**1**	**1**	**0**	**0**	**0**
9	**1**	**1**	**0**	**0**	**0**
2	0	0	0	0	0
3	1	1	1	1	1
4	1	0	1	0	0
5	**0**	**1**	**1**	**1**	**1**
7	**0**	**1**	**1**	**1**	**1**
8	**1**	**0**	**0**	**1**	**1**
10	**1**	**0**	**0**	**1**	**1**

The number of sites associated with a fingerprint (partition) can vary from one fingerprint to the other. In this example the partition associated with the [1 1 0 0 0] fingerprint is of size 3 as nucleotides 1, 6 and 9 share this finger print, while the partition associated with the [1 0 0 1 1] fingerprint is of size 2 (nucleotide 8 and 10).

III. Clean-sites protocol

In this protocol, we merged the predicted sites from all algorithms into overlapping regions (Supplementary Figure 12A). Each overlapping region is a union of predicted binding events from one or more algorithms. We then define a clean-site as the segment within the overlapping region that is predicted as part of a site by all algorithms that predicted a site anywhere in the overlapping region (see dark grey segment in the right region of Supplementary Figure 12B). We note that not all overlapping regions contain a clean-site, as shown in Supplementary Figure 12B where only the left region defines a clean-site. In our histone modification dataset, we only selected clean sites between 200bp and 500bp. Furthermore clean-sites less than 1000bp apart were also filtered out to avoid boundary effects since non-negligible amounts of di- and tri-nucleosomes are present in MNase-digested chromatin (see Supplementary Notes A.III). The main advantage of clean-sites is that they are not affected by boundary effects and can robustly distinguish between algorithms that predict an event in a region from those that do not. It should

be noted that the number of clean-sites is much smaller than the number of possible validation sites.

IV. Rationale behind the bias correction

The number of sites associated with a fingerprint (partition) can vary from one fingerprint to the other. (see above our fingerprint example). This means that if we were to estimate any quantity *Z* in a random fashion there will be a higher probability for selecting a site from a fingerprint (partition) with a large number of sites relative to a fingerprint with a smaller number of sites. Note that we evaluate the quantity *Z* within each fingerprint separately. To get an overall estimate of this quantity we have to compute its weighted sum across all fingerprints.

In an extreme example to demonstrate this point we choose two fingerprints: one fingerprint with one site and an estimated quantity Z=1 and another fingerprint which represents 99 sites and takes the value Z=0. We seek to estimate the mean of Z, which is 0.01. We can try the following 3 approaches:

1) We take a few random sites and estimate the average. We will most likely conclude that Z=0. The estimate will converge to the correct value when more sites are sampled.

2) We take the average of these two fingerprints and we will conclude that Z=0.5

3) We take the weighted average of the two fingerprints and we conclude that Z=0.01, which is equal to the true mean.

Supplementary Note D. – Visual comparison of overwhelmingly large numbers of algorithmic models

I. General considerations

Exploration of the parameter space of one algorithm easily leads to a large collection of predictions, each obtained by different parametric configurations. For instance, for an algorithm with k parameters, exhaustive exploration of all parametric configurations of n predefined values for each parameter will generate n^k predictions. Such number becomes prohibitive when exploring 15 algorithms, each with an average of 5 parameters for 7 possible values spanning a total of 4 orders of magnitude. To limit the size of the search space, we explored only predictions obtained varying one parameter at a time. In our study, this led to 315 models for 15 algorithms.

II. PCA

We used a principal component analysis to display quantitative profiles (e.g. number of predicted sites, performance, etc.) of the algorithmic parametric models (rows of the data matrix) and the datasets (columns of the data matrix) simultaneously as points in 2D space. This approach effectively visualizes similarities and dissimilarities between all the models. In addition, to better characterize the effects of each parameter, where needed, we added trajectories to the PCA plots by connecting those models where the same parameter was explored. For each algorithm, trajectories group parametric models where the same parameter was explored and connect incremental steps in the exploration of the same parameter. The model with default setting is, by definition, common to all trajectories. We also provided a series of background density maps, density isoclines for each algorithm and an optimized color-coding.

III. Bi-plot

The biplot provides an optimal approximation of the data matrix by such a 2D structure, in the sense that it displays the singular value decomposition, which gives the rank- two approximations to the data matrix having the smallest mean-squared error. The performance of a given algorithmic model in a dataset is approximated by the projection of the performance vector onto the direction of the dataset vector, multiplied by the length of the dataset vector. Thus, in this rank-two approximation, for a given algorithmic model and for dataset vectors of a given length, the algorithmic model has a high (or low) performance in datasets whose vector points in nearly the same (or opposite) direction as the algorithmic model.

Supplementary Table 1. Information about ChIP-Seq algorithms used in this study

a. List of the considered algorithms including version number, source location, required file format, and the programming language the given algorithm was developed in.

Program	*Version*	*Source location*	*Required format*	*Language*
CCAT	2.0	http://cmb.gis.a-star.edu.sg/ChIPSeq/tools.htm	BED	C
ChIPDiff		http://cmb.gis.a-star.edu.sg/ChIPSeq/paperChIPDiff.htm	TSV	C
ERANGE	3.2.1	http://woldlab.caltech.edu/rnaseq	BED	Python
FindPeaks	4.0.13	http://vancouvershortr.sourceforge.net/	BED	Java
FSeq	1.82	http://www.genome.duke.edu/labs/furey/software/fseq/	BED	Java
MACS	1.4	http://liulab.dfci.harvard.edu/MACS/	BED	Python
PeakSeq	1.01	http://archive.gersteinlab.org/proj/PeakSeq/	ELAND	Perl / C
Qeseq	0.2	http://sourceforge.net/projects/klugerlab/files/qeseq/v0.2.2/	TSV	C++
QuEST	2.4	http://mendel.stanford.edu/sidowlab/downloads/quest/	SSV	Perl
RSEG	0.0.0	http://smithlab.cmb.usc.edu/smithlab/index.php/software/rseg	BED	C++
SICER	1.03	http://home.gwu.edu/~wpeng/Software.htm	BED	Python
SISSRS	1.4	http://www.rajajothi.com/sissrs/	BED	Perl
SWEMBL	3.3.1	http://www.ebi.ac.uk/~swilder/SWEMBL/	BED	C
TPIC		http://bio.math.berkeley.edu/tpic/	BED	Perl/R
W-ChIPeaks		http://motif.bmi.ohio-state.edu/W-ChIPeaks	BED	Webserver

b. Publication dates of algorithms

Algorithm	*Date of publication*
CCAT	5-Apr-10
ChIPDiff	15-Oct-08
ERANGE	1-Jul-08
FindPeaks	1-Aug-08
FSeq	1-Nov-08
MACS	17-Sep-08
PeakSeq	27-Jan-09
QuEST	1-Sep-08
RSEG	15-Mar-11
SICER	1-Aug-09
SISSRS	1-Sep-08
SWEMBL	1-May-10
TPIC	12-Jan-11
W-ChIPeaks	1-Feb-11

Supplementary Table 2. Additional measures of default algorithmic performance in histone modification datasets

MYO.H3k27me3. GM.VDA	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	0.9909	0.6488	0.9976	0.8978	0.7842
ChIPDiff	0.9100	0.6937	0.9725	0.8928	0.7873
ERANGE	0.9247	0.0078	0.9997	0.7160	0.0155
FindPeaks	0.8077	0.7576	0.9277	0.8791	0.7818
FSeq	0.9984	0.5280	0.9997	0.8647	0.6907
MACS	1.0000	0.3166	1.0000	0.8045	0.4809
PeakSeq	0.9973	0.3002	0.9997	0.7996	0.4615
Qeseq	0.9816	0.6807	0.9949	0.9050	0.8039
QuEST	1.0000	0.0022	1.0000	0.7146	0.0045
RSEG	0.8023	0.7968	0.9214	0.8857	0.7996
SICER	-------*	0.0000	1.0000	0.7140	-------*
SISSRS	0.9661	0.0181	0.9997	0.7190	0.0355
SWEMBL	0.9888	0.6277	0.9972	0.8915	0.7679
TPIC	0.9935	0.6066	0.9984	0.8863	0.7533
W-ChIPeaks	1.0000	0.0103	1.0000	0.7169	0.0204
MYO.H3k36me3. GM.VDA	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	1.0000	0.2445	1.0000	0.5513	0.3930
ChIPDiff	0.8593	0.4459	0.8932	0.6275	0.5872
ERANGE	-------*	0.0000	1.0000	0.4060	-------*
FindPeaks	0.5742	0.9185	0.0035	0.5470	0.7066
FSeq	0.9850	0.2332	0.9948	0.5424	0.3771
MACS	1.0000	0.2295	1.0000	0.5424	0.3734
PeakSeq	0.9906	0.0979	0.9986	0.4636	0.1782
Qeseq	1.0000	0.3443	1.0000	0.6106	0.5123
QuEST	-------*	0.0000	1.0000	0.4060	-------*
RSEG	1.0000	0.2435	1.0000	0.5506	0.3916
SICER	1.0000	0.2424	1.0000	0.5500	0.3902
SISSRS	1.0000	0.0004	1.0000	0.4062	0.0007
SWEMBL	0.9909	0.2103	0.9972	0.5298	0.3469
TPIC	0.9892	0.2439	0.9961	0.5493	0.3913
W-ChIPeaks	-------*	0.0000	1.0000	0.4060	-------*

* Since the algorithm did not predict any of the validated sites as positive, precision and F-measure are not defined.

MYO.H3k27me3.GM	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	0.9554	0.6084	0.9920	0.9072	0.7434
ChIPDiff	0.8962	0.5954	0.9805	0.8954	0.7155
ERANGE	1.0000	0.0006	1.0000	0.7792	0.0011
FindPeaks	0.9399	0.6084	0.9890	0.9049	0.7387
FSeq	0.9682	0.5917	0.9945	0.9055	0.7345
MACS	1.0000	0.1959	1.0000	0.8224	0.3276
PeakSeq	1.0000	0.2457	1.0000	0.8334	0.3945
Qeseq	0.9440	0.6084	0.9898	0.9055	0.7399
QuEST	1.0000	0.0001	1.0000	0.7791	0.0002
RSEG	0.9440	0.6084	0.9898	0.9055	0.7399
SICER	-------*	0.0000	1.0000	0.7791	-------*
SISSRS	1.0000	0.0020	1.0000	0.7796	0.0041
SWEMBL	0.9512	0.6084	0.9911	0.9066	0.7421
TPIC	0.9459	0.6084	0.9901	0.9058	0.7405
W-ChIPeaks	1.0000	0.0092	1.0000	0.7811	0.0182

MYO.H3k27me3.MT	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	1.0000	0.4711	1.0000	0.9076	0.6404
ChIPDiff	0.9037	0.4788	0.9892	0.9000	0.6259
ERANGE	-------*	0.0000	1.0000	0.8253	-------*
FindPeaks	0.9906	0.4709	0.9991	0.9068	0.6384
FSeq	1.0000	0.4257	1.0000	0.8997	0.5972
MACS	1.0000	0.0788	1.0000	0.8391	0.1462
PeakSeq	1.0000	0.1181	1.0000	0.8459	0.2112
Qeseq	1.0000	0.4637	1.0000	0.9063	0.6336
QuEST	-------*	0.0000	1.0000	0.8253	-------*
RSEG	1.0000	0.4708	1.0000	0.9075	0.6402
SICER	-------*	0.0000	1.0000	0.8253	-------*
SISSRS	1.0000	0.0003	1.0000	0.8253	0.0006
SWEMBL	0.9831	0.4742	0.9983	0.9067	0.6398
TPIC	1.0000	0.4711	1.0000	0.9076	0.6404
W-ChIPeaks	1.0000	0.0003	1.0000	0.8258	0.0060

* Since the algorithm did not predict any of the validated sites as positive, precision and F-measure are not defined.

ES.H3K4me3	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	1.0000	0.6826	1.0000	0.9987	0.8114
ChIPDiff	0.0139	0.9683	0.7170	0.7181	0.0273
ERANGE	0.9964	0.6177	1.0000	0.9984	0.7626
FindPeaks	1.0000	0.7476	1.0000	0.9990	0.8556
FSeq	0.7641	0.9496	0.9988	0.9986	0.8468
MACS	0.8497	0.9267	0.9993	0.9990	0.8865
PeakSeq	0.0012	0.0039	0.9863	0.9823	0.0018
Qeseq	0.8577	0.9886	0.9993	0.9993	0.9185
QuEST	1.0000	0.3393	1.0000	0.9973	0.5067
RSEG	0.7791	0.9955	0.9988	0.9988	0.8741
SICER	0.8585	0.9946	0.9993	0.9993	0.9215
SISSRS	-------*	0.0000	1.0000	0.9959	-------*
SWEMBL	0.8318	0.8070	0.9993	0.9985	0.8192
TPIC	0.8811	0.9633	0.9995	0.9993	0.9203
W-ChIPeaks	1.0000	0.6746	1.0000	0.9987	0.8057

ES.H3K27me3	*Precision*	*Recall*	*True Negative Rate*	*Accuracy*	*F-measure*
CCAT	1.0000	0.0054	1.0000	0.9979	0.0107
ChIPDiff	0.0167	0.9557	0.8786	0.8788	0.0329
ERANGE	1.0000	0.1652	1.0000	0.9982	0.2835
FindPeaks	1.0000	1.0000	1.0000	1.0000	1.0000
FSeq	0.7110	0.6738	0.9994	0.9987	0.6919
MACS	1.0000	0.3515	1.0000	0.9986	0.5201
PeakSeq	1.0000	0.4588	1.0000	0.9988	0.6290
Qeseq	1.0000	0.7406	1.0000	0.9994	0.8509
QuEST	1.0000	0.1102	1.0000	0.9981	0.1985
RSEG	0.0194	1.0000	0.8908	0.8910	0.0381
SICER	0.7789	0.7710	0.9995	0.9990	0.7749
SISSRS	-------*	0.0000	1.0000	0.9978	-------*
SWEMBL	1.0000	0.7495	1.0000	0.9995	0.8568
TPIC	0.6215	0.7223	0.9990	0.9985	0.6681
W-ChIPeaks	1.0000	0.4370	1.0000	0.9988	0.6082

* Since the algorithm did not predict any of the validated sites as positive, precision and F-measure are not defined.

Supplementary Table 3. Number and average length of detected binding sites in 4 histone modification datasets

a. Average length of detected binding sites predicted by 15 algorithms in 4 existing qPCR validated ChIP-Seq datasets.

Algorithm	ES.H3K27me3	ES.H3K4me3	MYO.H3K27me3.GM	MYO.H3K36me3.GM
CCAT	929	1,438	3,400	2,895
ChIPDiff	6,198	4,961	36,599	25,780
ERANGE	803	1,069	2,334	1,426
FindPeaks	1,401	3,455	5,582	3,293
FSeq	163	303	376	349
MACS	1,325	1,859	1,058	890
PeakSeq	306	393	393	328
Qeseq	1,050	1,299	2,503	1,920
QuEST	1,659	897	5,294	4
RSEG	4,005	2,468	3,741	2,888
SICER	4,837	1,586	0	8,220
SISSRS	0	0	98	92
SWEMBL	983	682	4,839	3,222
TPIC	120	814	1,285	982
W-ChIPeaks	800	601	710	508

b. Number of detected binding sites predicted by 15 algorithms in 4 existing qPCR validated ChIP-Seq datasets.

Algorithm	ES.H3K27me3	ES.H3K4me3	MYO.H3K27me3.GM	MYO.H3K36me3.GM
CCAT	116	10,645	113,649	82,747
ChIPDiff	9,457	28,783	12,219	8,336
ERANGE	2,329	15,277	2,194	313
FindPeaks	80,227	7,124	106,223	215,313
FSeq	348,408	202,493	572,984	537,643
MACS	11,712	20,694	47,806	52,133
PeakSeq	37,024	256,823	256,823	234,352
Qeseq	46,213	52,212	159,091	182,981
QuEST	4,709	33,302	20	3
RSEG	19,503	36,019	29,941	11,406
SICER	27,814	40,719	0	36,248
SISSRS	0	0	30,596	12,840
SWEMBL	39,073	85,721	86,487	85,478
TPIC	580,523	62,613	258,519	269,257
W-ChIPeaks	3,167	4,329	3,131	2,381

Supplementary Table 4. List of all parametric models

Algorithm	*Parameters*	*Parameter values*	*H3K27 GM.VDA*	*H3K36 GM.VDA*	*H3K27 .GM*	*H3K27 .MT*	*ES.H3K4*	*ES.H3K27*
CCAT	default	NA	0.8232	0.6223	0.8002	0.7355	0.8413	0.5027
CCAT	BS	100	0.6541	0.5254	0.7814	0.7290	0.6875	0.5014
CCAT	BS	10	0.6302	0.5249	0.7754	0.7290	0.6827	0.5108
CCAT	BS	200	0.6542	0.5249	0.7813	0.7290	0.6826	0.5014
CCAT	BS	20	0.7526	0.5249	0.7813	0.7290	0.6875	0.5155
CCAT	BS	500	0.6542	0.5254	0.7813	0.7226	0.6826	0.5099
CCAT	BS	50	0.7527	0.5249	0.7813	0.7290	0.6875	0.5099
CCAT	BS	5	0.6541	0.5286	0.7734	0.7226	0.6703	0.5006
CCAT	MC	10	0.6542	0.5254	0.7813	0.7226	0.6826	0.5014
CCAT	MC	1	0.7525	0.5286	0.7803	0.7226	0.6875	0.5147
CCAT	MC	20	0.7460	0.5253	0.7822	0.7226	0.6626	0.5004
CCAT	MC	2	0.7525	0.5286	0.7803	0.7226	0.6875	0.5147
CCAT	MC	4	0.7527	0.5249	0.7813	0.7290	0.6875	0.5099
CCAT	MC	5	0.7527	0.5254	0.7814	0.7290	0.6826	0.5014
CCAT	MS	100	0.7438	0.5203	0.7713	0.7184	0.6612	0.5011
CCAT	MS	10	0.7478	0.5379	0.7825	0.7354	0.6546	0.5004
CCAT	MS	200	0.7545	0.5449	0.7796	0.7280	0.6620	0.5202
CCAT	MS	20	0.7466	0.5309	0.7834	0.7309	0.6601	0.5007
CCAT	MS	500	0.6713	0.5047	0.7153	0.6788	0.6277	0.5007
CCAT	MS	50	0.7527	0.5249	0.7813	0.7290	0.6875	0.5099
CCAT	MS	5	0.7443	0.5268	0.7762	0.7226	0.6614	0.5009
CCAT	MSc	10	0.5008	0.5000	0.5007	0.5000	0.5000	0.5000
CCAT	MSc	1	0.6150	0.6432	0.7987	0.7154	0.8151	0.6556
CCAT	MSc	2	0.5611	0.5386	0.7991	0.7355	0.7097	0.5947
CCAT	MSc	3	0.7527	0.5249	0.7813	0.7290	0.6875	0.5099
CCAT	MSc	5	0.6642	0.5000	0.6963	0.5633	0.6490	0.5000
CCAT	SW	1000	0.7054	0.6484	0.7830	0.7261	0.7574	0.5672
CCAT	SW	100	0.5923	0.5001	0.6724	0.5617	0.5713	0.5000
CCAT	SW	2000	0.8787	0.6849	0.7861	0.7249	0.7118	0.6170
CCAT	SW	200	0.6898	0.5006	0.7640	0.6351	0.6250	0.5000
CCAT	SW	5000	0.6539	0.6794	0.8562	0.8191	0.6917	0.6479
CCAT	SW	500	0.7527	0.5249	0.7813	0.729	0.6875	0.5099
CCAT	SW	50	0.5043	0.5	0.5041	0.5001	0.5	0.5
ChIPDiff	default	NA	0.8331	0.6696	0.7879	0.734	0.8426	0.9172
ChIPDiff	MAXIT	1000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	100	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	2000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	200	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	5000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	500	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXIT	50	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXTR	100000	0.6515	0.5011	0.7033	0.6216	0.8887	0.8855
ChIPDiff	MAXTR	10000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXTR	1000	0.6591	0.5011	0.7027	0.6854	0.8887	0.8855

ChIPDiff	MAXTR	100	0.6591	0.5	0.7027	0.69	0.9106	0.8855
ChIPDiff	MAXTR	20000	0.6591	0.5011	0.7027	0.6854	0.8887	0.8855
ChIPDiff	MAXTR	2000	0.6591	0.5011	0.7027	0.6854	0.8887	0.8855
ChIPDiff	MAXTR	200	0.6591	0.5011	0.7027	0.6851	0.9106	0.8855
ChIPDiff	MAXTR	50000	0.6515	0.5011	0.7033	0.6854	0.8887	0.8855
ChIPDiff	MAXTR	5000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MAXTR	500	0.6591	0.5011	0.7027	0.6854	0.8887	0.8855
ChIPDiff	MINFOLD	10	0.5076	0.5	0.5286	0.5086	0.8196	0.8855
ChIPDiff	MINFOLD	20	0.5008	0.5	0.5036	0.5	0.8132	0.8855
ChIPDiff	MINFOLD	2	0.7034	0.5295	0.7683	0.7303	0.972	0.8719
ChIPDiff	MINFOLD	3	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MINFOLD	5	0.614	0.5	0.633	0.5938	0.9106	0.8855
ChIPDiff	MINP	50	0.679	0.5016	0.7389	0.7021	0.9342	0.8235
ChIPDiff	MINP	80	0.6598	0.5016	0.7278	0.69	0.9338	0.8719
ChIPDiff	MINP	90	0.6591	0.5011	0.7219	0.69	0.8887	0.8855
ChIPDiff	MINP	95	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MINP	98	0.6524	0.5	0.6921	0.6747	0.9106	0.8855
ChIPDiff	MINP	99	0.6524	0.5	0.6894	0.6687	0.8828	0.8855
ChIPDiff	MINR	10000	0.6515	0.5261	0.7089	0.6436	0.9111	0.8371
ChIPDiff	MINR	1000	0.6591	0.5011	0.7027	0.6854	0.9106	0.8855
ChIPDiff	MINR	100	0.6586	0.5011	0.7056	0.6738	0.9106	0.802
ChIPDiff	MINR	2000	0.6591	0.5	0.7027	0.6618	0.9106	0.8989
ChIPDiff	MINR	200	0.6586	0.5011	0.7056	0.6738	0.9106	0.802
ChIPDiff	MINR	5000	0.6536	0.5261	0.7089	0.6597	0.8893	0.8989
ChIPDiff	MINR	500	0.6586	0.5011	0.7056	0.6738	0.9106	0.802
W-ChIPeaks	default	NA	0.5085	0.5000	0.5046	0.5015	0.8373	0.7185
W-ChIPeaks	P	0.98	0.5052	0.5000	0.5018	0.5006	0.7791	0.6688
W-ChIPeaks	P	0.995	0.5011	0.5000	0.5006	0.5005	0.6819	0.6038
W-ChIPeaks	P	0.999	0.5002	0.5000	0.5000	0.5005	0.5274	0.5301
W-ChIPeaks	P	0.99	0.5025	0.5000	0.5007	0.5005	0.7399	0.6153
ERANGE	default	NA	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	autoshift	0	0.5247	0.5	0.5076	0.5	0.8099	0.5826
ERANGE	autoshift	1	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	100	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	10	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	1	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	20	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	2	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	50	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	min	5	0.5038	0.5	0.5003	0.5	0.8088	0.5826
ERANGE	nodirect	0	0.5046	0.5	0.501	0.5005	0.8216	0.6261
ERANGE	notrim	0	0.5038	0.5	0.5003	0.5	0.8097	0.5826
ERANGE	ratio	100	0.5038	0.5	0.5003	0.5	0.5	0.5
ERANGE	ratio	10	0.5038	0.5	0.5003	0.5	0.7701	0.5216
ERANGE	ratio	1	0.5038	0.5	0.5003	0.5	0.8088	0.6076
ERANGE	ratio	20	0.5038	0.5	0.5003	0.5	0.7281	0.501
ERANGE	ratio	2	0.5038	0.5	0.5003	0.5	0.8088	0.6076
ERANGE	ratio	50	0.5038	0.5	0.5003	0.5	0.5	0.5
ERANGE	ratio	5	0.5038	0.5	0.5003	0.5	0.8088	0.5826

ERANGE	shift	1000	0.5	0.5	0.5	0.5	0.5	0.5
ERANGE	shift	100	0.5029	0.5	0.5069	0.5005	0.7988	0.6062
ERANGE	shift	2000	0.5	0.5	0.5	0.5	0.5	0.5
ERANGE	shift	200	0.5046	0.5	0.5013	0.5005	0.6329	0.5539
ERANGE	shift	20	0.5236	0.5	0.5018	0.5005	0.8102	0.6242
ERANGE	shift	500	0.5224	0.5	0.5018	0.5	0.5	0.5
ERANGE	shift	50	0.5237	0.5	0.5081	0.5005	0.8022	0.6244
ERANGE	space	1000	0.5671	0.4999	0.566	0.5084	0.7352	0.5198
ERANGE	space	100	0.5059	0.5	0.5189	0.5005	0.8591	0.6455
ERANGE	space	2000	0.4859	0.5845	0.5283	0.5141	0.6348	0.6245
ERANGE	space	200	0.5476	0.5	0.527	0.5084	0.877	0.8016
ERANGE	space	5000	0.4889	0.611	0.5189	0.5557	0.7268	0.5867
ERANGE	space	500	0.5524	0.4997	0.5712	0.5153	0.8769	0.8649
ERANGE	space	50	0.5038	0.5	0.5003	0.5	0.8088	0.5826
FindPeaks	default	NA	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	disttype	0	0.7427	0.6211	0.7825	0.6663	0.8555	0.8894
FindPeaks	it	100	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	it	10	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	it	1	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	it	20	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	it	50	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	it	5	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	logtransform	0	0.6161	0.4659	0.7921	0.6604	0.5217	0.5
FindPeaks	mincove	2	0.8469	0.4553	0.7987	0.735	0.8752	1
FindPeaks	mincove	3	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	mincove	4	0.8427	0.4629	0.7987	0.735	0.8738	1
FindPeaks	mincove	5	0.7444	0.4629	0.7987	0.735	0.8738	1
FindPeaks	mincove	6	0.7444	0.4629	0.7987	0.735	0.8738	1
FindPeaks	subpeaks	0.01	0.8467	0.6481	0.7987	0.7241	0.8791	1
FindPeaks	subpeaks	0.02	0.8254	0.7334	0.7987	0.7192	0.823	1
FindPeaks	subpeaks	0.05	0.7717	0.7576	0.7991	0.7102	0.823	1
FindPeaks	subpeaks	0.1	0.7678	0.6823	0.7839	0.6987	0.823	0.9975
FindPeaks	subpeaks	0.2	0.7198	0.6956	0.769	0.6797	0.8196	0.9975
FindPeaks	subpeaks	0.5	0.6145	0.6174	0.7125	0.6064	0.8428	0.8993
FindPeaks	subpeaks	0.8	0.5983	0.6143	0.6658	0.5915	0.812	0.8733
FindPeaks	subpeaks	0.9	0.6082	0.5553	0.6454	0.5709	0.8033	0.8733
FindPeaks	trim	0.01	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.02	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.05	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.1	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.2	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.5	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.8	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	trim	0.9	0.8427	0.461	0.7987	0.735	0.8738	1
FindPeaks	window	100	0.7727	0.6961	0.7987	0.8386	0.8873	0.9986
FindPeaks	window	20	0.7369	0.7075	0.7987	0.8423	0.8842	0.9938
FindPeaks	window	50	0.7669	0.7005	0.7987	0.8423	0.8873	0.9933
FSeq	default	NA	0.7638	0.614	0.7931	0.7128	0.9742	0.8366
FSeq	l	1000	0.7604	0.614	0.7827	0.7221	0.9797	0.844

FSeq	l	100	0.7406	0.5691	0.6942	0.6165	0.9072	0.7589
FSeq	l	2000	0.7532	0.5938	0.7781	0.7289	0.9951	0.8163
FSeq	l	200	0.7847	0.582	0.751	0.6683	0.9296	0.7962
FSeq	l	5000	0.718	0.5021	0.7933	0.7272	0.9995	0.802
FSeq	l	500	0.8421	0.614	0.7809	0.7048	0.9715	0.8273
FSeq	l	50	0.6838	0.5487	0.6333	0.5767	0.8765	0.72
FSeq	t	10	0.6358	0.5271	0.6407	0.5814	0.922	0.7306
FSeq	t	1	0.5536	0.6355	0.799	0.7752	0.8713	0.8627
FSeq	t	2	0.6892	0.619	0.7998	0.7299	0.9204	0.8831
FSeq	t	5	0.7576	0.6106	0.7653	0.6959	0.9653	0.8219
MACS	default	NA	0.6583	0.6148	0.5980	0.5394	0.963	0.6757
MACS	mfold	10.2	0.5	0.6236	0.5	0.5063	0.9562	0.8431
MACS	mfold	1.5	0.5	0.6246	0.5	0.5553	0.9689	0.8328
MACS	mfold	20.5	0.5502	0.5954	0.5201	0.5063	0.9035	0.8411
MACS	mfold	50.1	0.5502	0.5954	0.5201	0.5063	0.9697	0.8377
MACS	mfold	5.1	0.5	0.6247	0.5	0.5509	0.902	0.8452
MACS	nolambda	0	0.5502	0.5954	0.5201	0.5063	0.9697	0.8377
MACS	nolambda	1	0.5502	0.5954	0.5201	0.5063	0.9562	0.8377
MACS	pval	0.1	0.5582	0.6236	0.5476	0.5063	0.9689	0.8401
MACS	pval	0.01	0.6736	0.6186	0.5559	0.5166	0.9875	0.864
MACS	pval	0.001	0.6668	0.6159	0.584	0.5208	0.9886	0.8674
MACS	pval	0.0001	0.6591	0.6154	0.5936	0.5338	0.963	0.8592
MACS	pval (default)	0.00001	0.6583	0.6148	0.598	0.5394	0.963	0.6757
MACS	pval	0.000001	0.6127	0.5917	0.5862	0.5468	0.9401	0.6757
MACS	pval	0.0000001	0.5268	0.5658	0.5717	0.5365	0.9231	0.6757
MACS	pval	0.00000001	0.5922	0.5914	0.6061	0.5354	0.8283	0.6757
PeakSeq	default	NA	0.65	0.5483	0.6229	0.559	0.4951	0.7294
PeakSeq	bin	100000	0.6339	0.5269	0.674	0.5628	0.9264	0.7281
PeakSeq	bin	1000	0.6498	0.549	0.6855	0.572	0.9264	0.7302
PeakSeq	bin	20000	0.6497	0.5486	0.6683	0.558	0.9289	0.7288
PeakSeq	bin	2000	0.6498	0.5509	0.6846	0.5673	0.9264	0.7294
PeakSeq	bin	50000	0.6339	0.5486	0.6683	0.5628	0.9264	0.7289
PeakSeq	bin	5000	0.6498	0.5486	0.6851	0.5629	0.9264	0.7366
PeakSeq	fdr	0.1	0.6506	0.5487	0.7022	0.5783	0.929	0.7583
PeakSeq	fdr	0.01	0.6496	0.5185	0.6347	0.5462	0.9233	0.7001
PeakSeq	fdr	0.001	0.5864	0.505	0.6109	0.5283	0.9192	0.6924
PeakSeq	fdr	0.0001	0.5996	0.505	0.6097	0.5267	0.9224	0.6908
PeakSeq	fdr	0.00001	0.5898	0.5047	0.5988	0.5221	0.9204	0.6915
PeakSeq	fdr	0.000001	0.5863	0.5008	0.6086	0.5221	0.9079	0.6913
PeakSeq	fdr	0.0000001	0.5741	0.5048	0.5993	0.5262	0.9198	0.6881
PeakSeq	fdr	0.00000001	0.574	0.5032	0.6045	0.5193	0.9202	0.6966
PeakSeq	maxcount	100	0.6497	0.5487	0.6803	0.5628	0.9264	0.7281
PeakSeq	maxcount	10	0.6497	0.5486	0.6731	0.5629	0.9265	0.7281
PeakSeq	maxcount	20	0.6497	0.5486	0.6798	0.5629	0.9264	0.7281
PeakSeq	maxcount	50	0.6497	0.5468	0.6731	0.5633	0.9265	0.7281
PeakSeq	maxcount	5	0.6497	0.5468	0.6806	0.5628	0.9264	0.7281
PeakSeq	maxext	10000	0.6497	0.5486	0.6793	0.563	0.9264	0.7281
PeakSeq	maxext	100	0.6497	0.5468	0.6736	0.5582	0.9264	0.7353
PeakSeq	maxext	2000	0.6497	0.5486	0.6736	0.5634	0.9264	0.7281

PeakSeq	maxext	200	0.6497	0.5486	0.6731	0.5629	0.9264	0.7289
PeakSeq	maxext	5000	0.6497	0.5486	0.6726	0.5629	0.9264	0.728
PeakSeq	maxext	500	0.6497	0.5486	0.6736	0.5629	0.9264	0.728
PeakSeq	maxgap	1000	0.6497	0.5444	0.7117	0.6145	0.5	0.75
PeakSeq	maxgap	100	0.6253	0.5484	0.673	0.5635	0.9267	0.7083
PeakSeq	maxgap	2000	0.7064	0.5188	0.7289	0.6307	0.5	0.75
PeakSeq	maxgap	20	0.6263	0.5265	0.6557	0.5	0.5	0.5
PeakSeq	maxgap	500	0.6497	0.5468	0.6842	0.5908	0.5	0.7495
PeakSeq	maxgap	50	0.6261	0.5482	0.6666	0.5585	0.5	0.7083
PeakSeq	nsims	100	0.6497	0.5486	0.6731	0.5629	0.9264	0.7281
PeakSeq	nsims	1	0.6497	0.5468	0.6806	0.5631	0.9289	0.736
PeakSeq	nsims	20	0.6497	0.5468	0.6736	0.5628	0.9264	0.7281
PeakSeq	nsims	2	0.6497	0.5486	0.6731	0.5633	0.9264	0.7275
PeakSeq	nsims	50	0.6497	0.5486	0.6793	0.5628	0.9264	0.728
PeakSeq	nsims	5	0.6497	0.5468	0.6736	0.563	0.9264	0.7281
PeakSeq	pval	0.1	0.6498	0.5505	0.6851	0.5628	0.9265	0.7294
PeakSeq	pval	0.01	0.6339	0.5269	0.6619	0.5466	0.9262	0.7266
PeakSeq	pval	0.001	0.614	0.5005	0.6366	0.5466	0.9287	0.7259
PeakSeq	pval	0.0001	0.6139	0.5	0.6126	0.5413	0.9262	0.7014
PeakSeq	pval	0.00001	0.6139	0.5	0.6118	0.5333	0.9287	0.7006
PeakSeq	pval	0.000001	0.5907	0.5	0.6051	0.5294	0.8804	0.7006
PeakSeq	pval	0.0000001	0.5905	0.5	0.6051	0.5257	0.8803	0.6516
PeakSeq	pval	0.00000001	0.5902	0.5	0.6055	0.5205	0.8803	0.6524
PeakSeq	region	10000	0.6497	0.5486	0.6731	0.5629	0.9287	0.7281
PeakSeq	region	1000	0.6497	0.5486	0.6731	0.5633	0.9264	0.728
PeakSeq	region	20000	0.6497	0.5468	0.6735	0.5634	0.9264	0.7274
PeakSeq	region	200	0.6497	0.5486	0.673	0.5629	0.9264	0.7281
PeakSeq	region	5000	0.6497	0.5486	0.6798	0.5627	0.9264	0.7281
PeakSeq	region	500	0.6497	0.5486	0.6798	0.5632	0.9264	0.7289
PeakSeq	wperc	1000	0.6497	0.5486	0.6731	0.5628	0.9265	0.7281
PeakSeq	wperc	100	0.5	0.5	0.5	0.5	0.5	0.5
PeakSeq	wperc	2000	0.6497	0.5486	0.6731	0.5632	0.5	0.7289
PeakSeq	wperc	200	0.6497	0.5486	0.6736	0.5628	0.9264	0.728
PeakSeq	wperc	20	0.5	0.5	0.5	0.5	0.5	0.5
PeakSeq	wperc	500	0.6497	0.5468	0.6731	0.5634	0.9264	0.7281
PeakSeq	wperc	50	0.5	0.5	0.5	0.5	0.5	0.5
PeakSeq	wsize	100000	0.5	0.5	0.5	0.5	0.5	0.5
PeakSeq	wsize	2000000	0.5511	0.5005	0.569	0.5164	0.9223	0.6905
PeakSeq	wsize	200000	0.5	0.5	0.5	0.5	0.5	0.5
PeakSeq	wsize	5000000	0.5017	0.5	0.504	0.5	0.842	0.6362
PeakSeq	wsize	500000	0.5	0.5	0.5	0.5	0.5	0.5
Qeseq	default	NA	0.8378	0.6722	0.7991	0.7319	0.994	0.8703
QuEST	default	NA	0.5011	0.5	0.5	0.5	0.6697	0.5551
RSEG	default	NA	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	bin	10000	0.5693	0.7516	0.7606	0.8382	0.7886	0.861
RSEG	bin	1000	0.7559	0.7418	0.7991	0.7568	0.9698	0.9454
RSEG	bin	2000	0.6714	0.7332	0.7991	0.7372	0.9701	0.9454
RSEG	bin	5000	0.5715	0.746	0.7976	0.8387	0.9661	0.861
RSEG	desert	1000	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454

RSEG	desert	100	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	desert	2000	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	desert	200	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	desert	5000	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	desert	500	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	Gauss	0	0.5	0.5	0.5	0.5	0.5	0.5
RSEG	iter	1000	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	iter	100	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	iter	2000	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	iter	200	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	iter	500	0.8591	0.6217	0.7991	0.7354	0.9972	0.9454
RSEG	p	0.01	0.8591	0.6218	0.7991	0.7355	0.9968	0.9454
RSEG	p	0.05	0.7704	0.6214	0.7991	0.7355	0.9676	0.9454
RSEG	p	0.1	0.7671	0.5353	0.7991	0.7355	0.9676	0.9469
RSEG	Poisson	0	0.663	0.6822	0.5	0.7311	0.9376	0.6844
SICER	default	NA	0.5	0.6212	0.5	0.5	0.997	0.8853
SICER	W	100	0.5	0.6019	0.5	0.7355	0.9925	0.8855
SICER	G	1000	0.5	0.6233	0.5	0.5	0.9676	0.8989
SICER	G	2000	0.5	0.7038	0.5	0.5	0.8841	0.9777
SICER	G	200	0.5	0.6201	0.5	0.5	0.9956	0.802
SICER	FDR	1	0.5	0.6212	0.5	0.5	0.997	0.8853
SICER	FDR	2	0.5	0.6212	0.5	0.5	0.997	0.8853
SICER	FDR	4	0.5	0.6212	0.5	0.5	0.997	0.8018
SICER	FDR	5	0.5	0.5927	0.5	0.5	0.997	0.8018
SICER	FDR	6	0.5	0.5926	0.5	0.5	0.997	0.8018
SICER	FDR	7	0.5	0.5926	0.5	0.5	0.997	0.8018
SICER	FDR	8	0.5	0.5926	0.5	0.5	0.997	0.8018
SICER	W	20	0.5	0.5738	0.5	0.7355	0.9118	0.8855
SICER	W	50	0.5	0.5307	0.5	0.7354	0.9742	0.8855
SISSRS	default	NA	0.5089	0.5002	0.501	0.5002	0.5	0.5
SISSRS	dirreads	10	0.5016	0.5	0.5018	0.5005	0.6822	0.606
SISSRS	dirreads	1	0.5366	0.5	0.5147	0.5006	0.6914	0.6329
SISSRS	dirreads	20	0.5	0.5	0.5	0.5	0.5973	0.5005
SISSRS	dirreads	2	0.5366	0.5	0.5147	0.5006	0.6914	0.6329
SISSRS	dirreads	5	0.5366	0.5	0.5147	0.5006	0.6901	0.6329
SISSRS	pval	0.1	0.5466	0.5094	0.5157	0.5036	0.6326	0.607
SISSRS	pval	0.01	0.5427	0.5094	0.5082	0.5002	0.5439	0.5008
SISSRS	pval	0.001	0.5089	0.5002	0.501	0.5002	0.5	0.5
SISSRS	pval	0.0001	0.502	0.5002	0.5	0.5	0.5	0.5
SISSRS	window	10	0.5403	0.5002	0.5236	0.5068	0.6814	0.6258
SISSRS	window	20	0.5466	0.5094	0.5157	0.5036	0.6326	0.607
SISSRS	window	2	0.541	0.5002	0.5269	0.5043	0.835	0.6531
SISSRS	window	6	0.5413	0.5055	0.5245	0.5047	0.8049	0.6533
SWEMBL	default	NA	0.8124	0.6037	0.7998	0.7362	0.9032	0.8747
SWEMBL	d	1000	0.7046	0.5491	0.7433	0.667	0.7711	0.6933
SWEMBL	d	100	0.6875	0.5491	0.7291	0.658	0.7711	0.6933
SWEMBL	d	200	0.7046	0.5491	0.7433	0.6623	0.7711	0.6933
SWEMBL	d	500	0.7046	0.5491	0.7355	0.667	0.7711	0.6933
SWEMBL	d	50	0.6952	0.5433	0.7308	0.6301	0.7611	0.6926

SWEMBL	f	150	0.8148	0.6019	0.7998	0.7362	0.9629	0.8783
SWEMBL	f	200	0.9128	0.619	0.7998	0.7362	0.9555	0.8855
SWEMBL	f	50	0.8137	0.6049	0.7998	0.7362	0.9741	0.8775
SWEMBL	p	0.01	0.8124	0.6037	0.7998	0.7362	0.9032	0.8747
SWEMBL	p	0.05	0.7067	0.5491	0.7506	0.6671	0.7711	0.6933
SWEMBL	p	0.1	0.5567	0.5085	0.5923	0.5198	0.7251	0.6049
SWEMBL	p	0.5	0.5003	0.5	0.5001	0.5	0.5097	0.5
TPIC	default	NA	0.8025	0.62	0.7993	0.7355	0.9814	0.8607
TPIC	D	10	0.779	0.6071	0.8023	0.723	0.9641	0.8333
TPIC	D	1	0.8799	0.6066	0.7902	0.7034	0.9641	0.8333
TPIC	D	20	0.7796	0.6071	0.8023	0.723	0.9641	0.8333
TPIC	D	2	0.7818	0.6069	0.7926	0.7063	0.9641	0.8333
TPIC	D	50	0.8776	0.6071	0.8023	0.723	0.9641	0.8333
TPIC	MRL	100000	0.7802	0.6072	0.8042	0.7093	0.9641	0.8333
TPIC	MRL	1000	0.779	0.6072	0.8023	0.723	0.9641	0.8333
TPIC	MRL	20000	0.779	0.6072	0.8042	0.7084	0.9641	0.8333
TPIC	MRL	2000	0.7793	0.6073	0.8023	0.723	0.9641	0.8333
TPIC	MRL	50000	0.7794	0.6071	0.8023	0.7084	0.9641	0.8333
TPIC	MRL	5000	0.7785	0.6071	0.8023	0.723	0.9641	0.8333
TPIC	W	10000	0.779	0.6069	0.8023	0.723	0.9641	0.8333
TPIC	W	100	0.8811	0.6073	0.795	0.7066	0.9641	0.8333
TPIC	W	2000	0.7785	0.6071	0.8023	0.723	0.9641	0.8333
TPIC	W	200	0.7832	0.6076	0.7976	0.7034	0.9641	0.8333
TPIC	W	5000	0.7789	0.6066	0.8042	0.723	0.9641	0.8333
TPIC	W	500	0.8797	0.6079	0.7973	0.711	0.9641	0.8333

Supplementary Table 5. Characterization of the transcription factor datasets used in this study

a. Transcription factor ChIP-Seq datasets in terms of the number of ChIP and control reads.

Transcription factor	*ChIP reads*	*Control reads*
MYO.E2F4.MT	10,141,092	16,845,907
MYO.SIN3A.MT	2,920,367	5,922,554
MYO.SIN3A.GM	8,208,294	9,798,009
MYO.SIN3B.MT	2,334,764	5,922,554
MYO.SIN3B.GM	4,563,905	9,798,009
STAT1	26,731,492	23,435,631

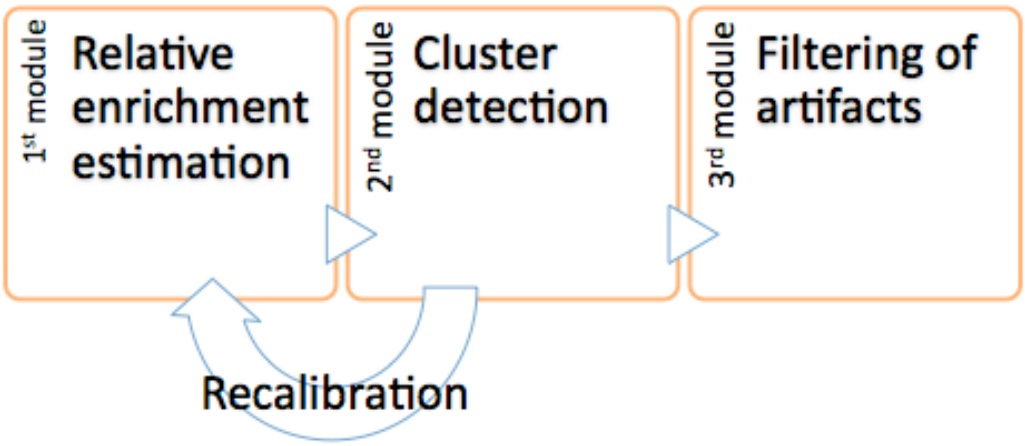
b. Characterization of the qPCR validation sites.

Transcription factor	*Positives*	*Negatives*	*Total*	*Source*
MYO.E2F4.MT	51	12	63	(34)
MYO.SIN3A.MT	4	61	65	(34)
MYO.SIN3A.GM	11	54	65	(34)
MYO.SIN3B.MT	16	49	65	(34)
MYO.SIN3B.GM	23	42	65	(34)
STAT1	120	160	280	(16)

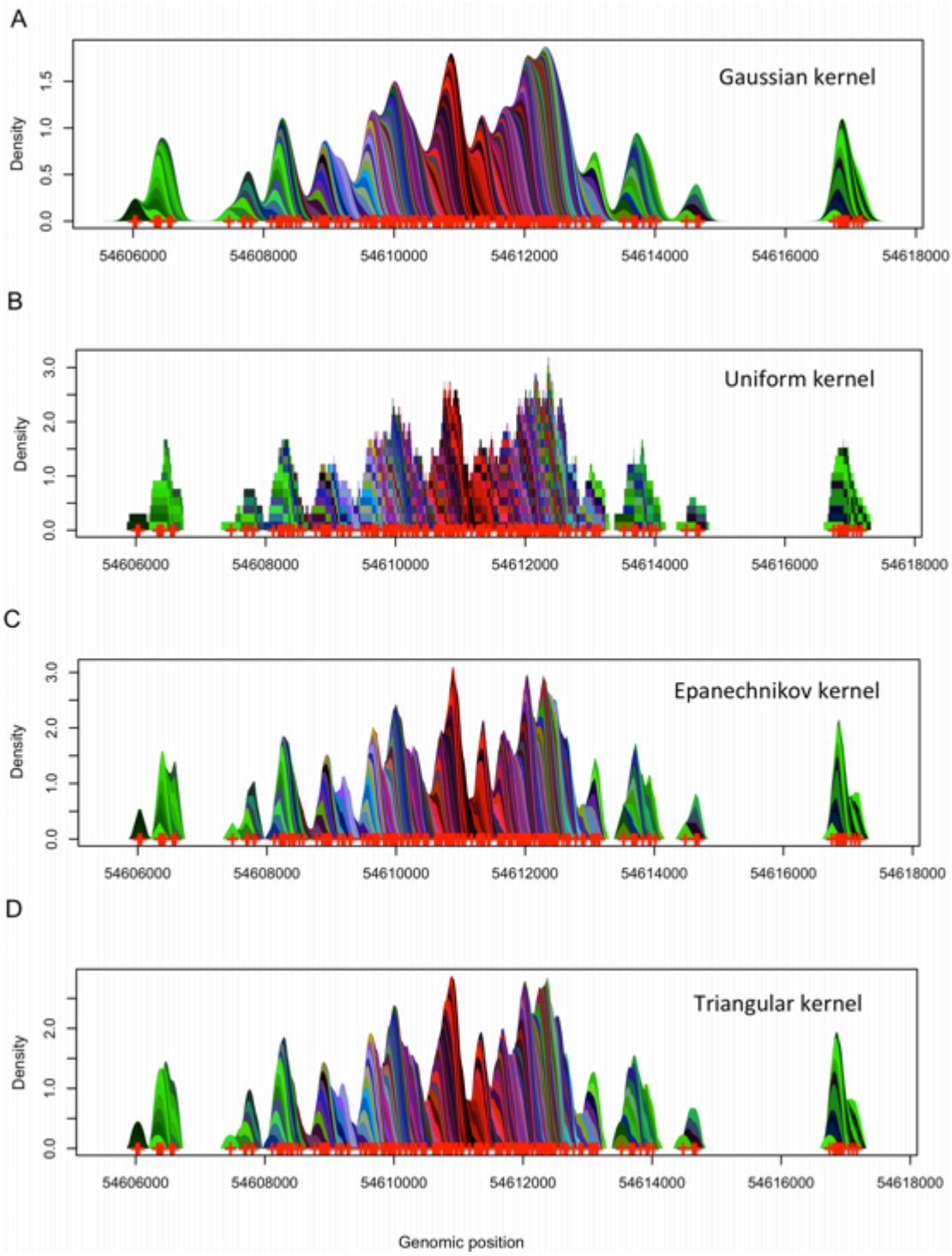
Supplementary Table 6. Performance analysis based on unbiased AUCROC statistic for each algorithm in transcription factor datasets

Algorithm	MYO.E2F4.MT	MYO.SIN3A.MT	MYO.SIN3A.GM	MYO.SIN3B.MT	MYO.SIN3B.GM	STAT1
CCAT	0.5001	0.8770	0.5000	0.5564	0.5011	0.5001
ChIPDiff	0.5750	0.7626	0.6389	0.6658	0.8007	0.5048
ERANGE	0.5001	0.5299	0.5000	0.5185	0.5003	0.5001
FindPeaks	0.4991	0.9696	0.9755	0.5448	0.5014	0.5001
FSeq	0.5010	0.9886	0.7806	0.5549	0.5001	0.5004
MACS	0.5001	0.8537	0.5000	0.5559	0.5012	0.5000
PeakSeq	0.5001	0.8771	0.5007	0.5522	0.5011	0.5001
Qeseq	0.5001	0.9998	0.5145	0.5547	0.5013	0.5000
QuEST	0.5000	0.5234	0.5000	0.5133	0.5001	0.5000
RSEG	0.4232	0.3370	0.4731	0.3648	0.4696	0.5001
SICER	0.5001	0.9989	0.5086	0.5554	0.5013	0.5004
SISSRS	0.5000	0.8342	0.5000	0.5355	0.5004	0.5000
SWEMBL	0.5001	0.5000	0.5026	0.5555	0.5012	0.5005
TPIC	0.5001	0.9873	0.5012	0.5541	0.5002	0.5000
W-ChIPeaks	0.5001	0.9063	0.5009	0.5513	0.5012	0.5001

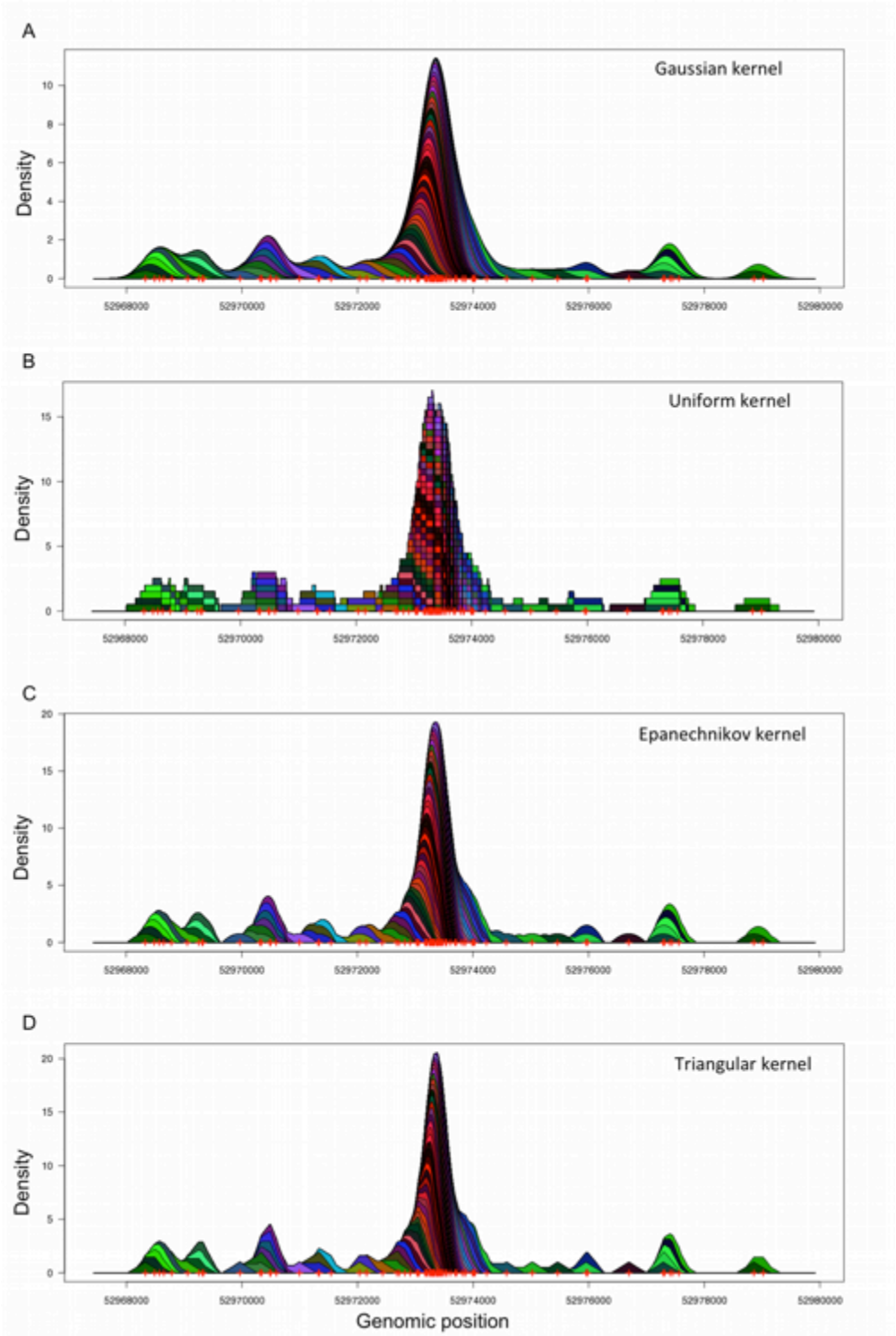
Supplementary Figures



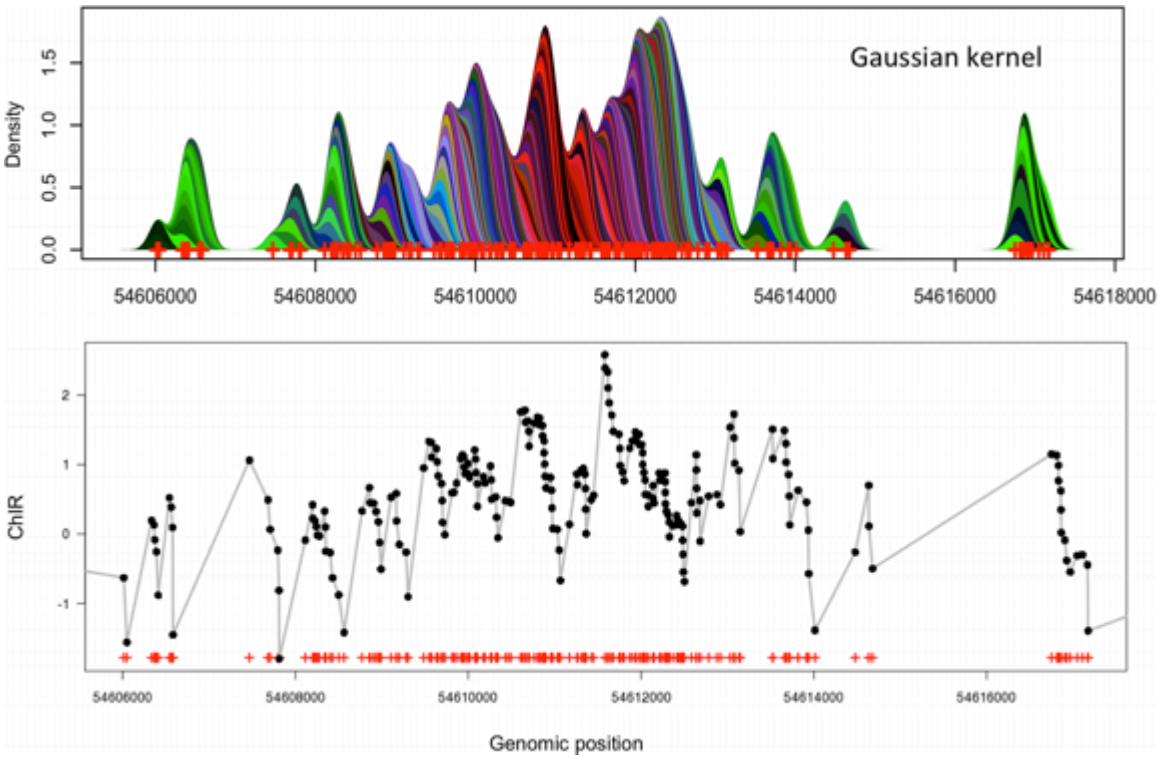
Supplementary Figure 1. Display of Qeseq's scheme. Qeseq consists of three main modules: relative enrichment estimation, cluster detection and filtering of artifacts. Following recalibration, iterative relative enrichment estimation and cluster detection are repeated until no new events are identified. Subsequently, artifact filtering is performed only once at the last step.



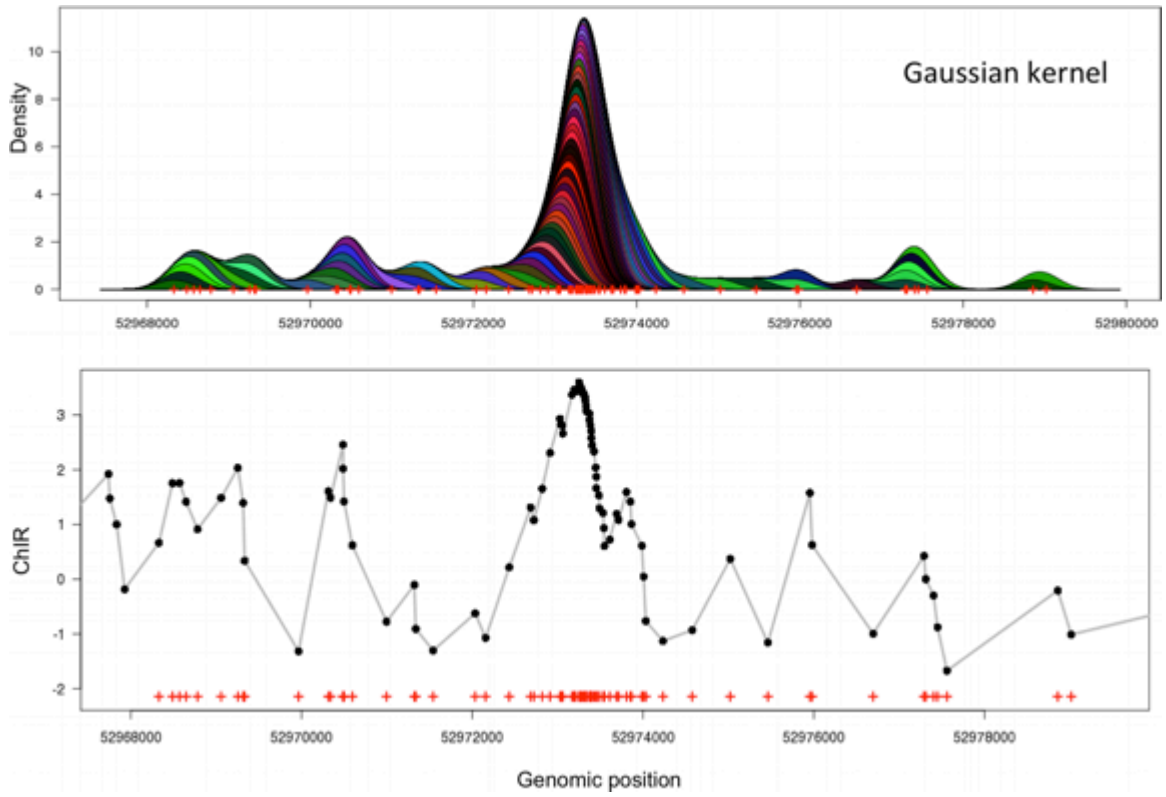
Supplementary Figure 2. Illustration of kernel density estimators in genomic region chr19: 54,605,999-54,617,335 of the MYO.H3K27me3.GM ChIP-Seq dataset. The genomic positions of ChIP reads are marked with red crosses in all panels. This figure shows that regardless of the choice of kernel, the derived signals are very similar. Densities have been linearly rescaled for visualization. **(A)** Density estimated using a uniform kernel with 150bp kernel bandwidth. **(B)** Density estimates using a Gaussian kernel (Equation 1, h=150bp). **(C)** Epanechnikov kernel density with 150bp kernel bandwidth. **(D)** Triangular kernel density with 150bp kernel bandwidth.



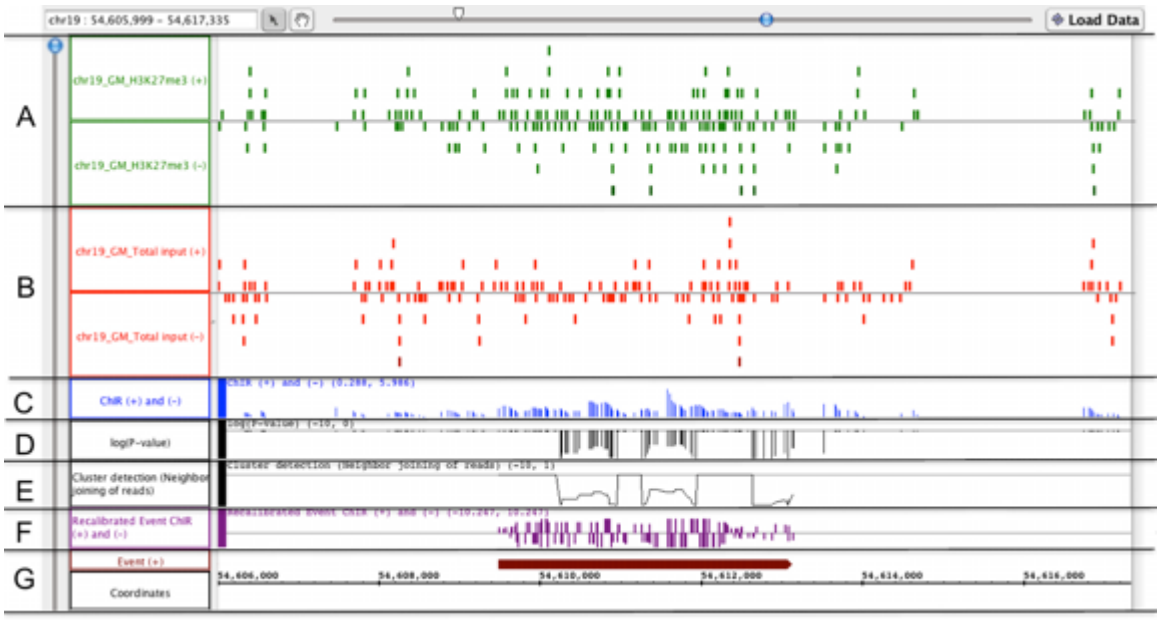
Supplementary Figure 3. Illustration of kernel density estimators in genomic region chr7:52,968,000-52,981,000 of the MYO.SIN3B.GM ChIP-Seq dataset. The genomic positions of ChIP reads are marked with red crosses in all panels. This figure shows that regardless of the choice of kernel, the derived signals are very similar. **(A)** Density estimated using a uniform kernel with 150bp kernel bandwidth. **(B)** Density estimates using a Gaussian kernel (Equation 1, h=150bp). **(C)** Epanechnikov kernel density with 150bp kernel bandwidth. **(D)** Triangular kernel density with 150bp kernel bandwidth.



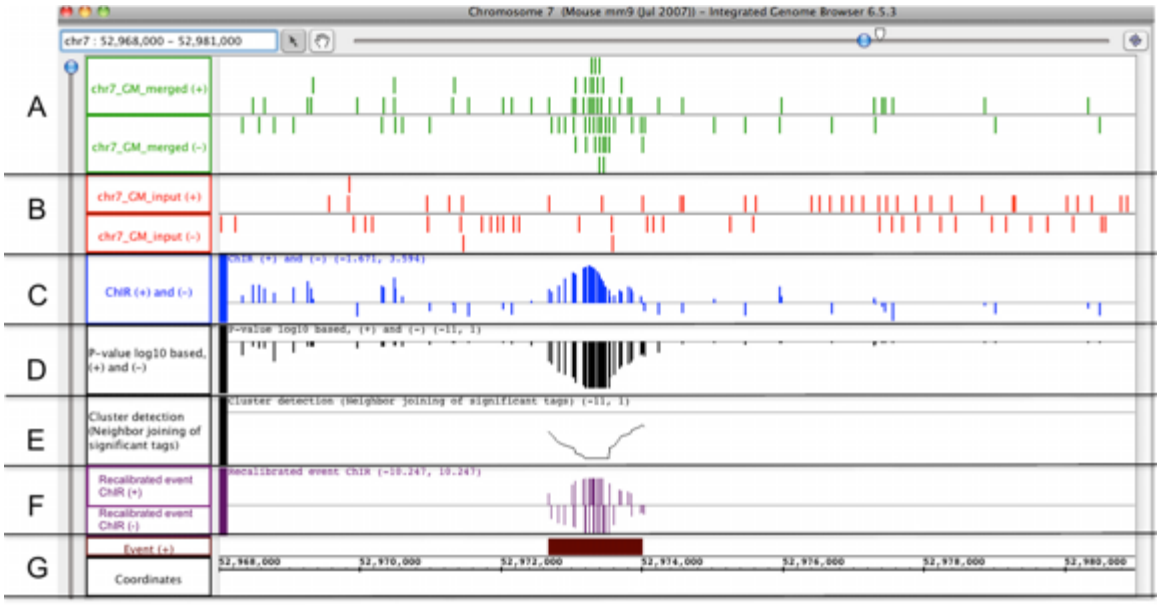
Supplementary Figure 4. Illustration of Gaussian kernel density and derived ChIR in genomic region chr19: 54,605,999-54,617,335 of the MYO.H3K27me3.GM ChIP-Seq dataset. The genomic positions of individual ChIP reads are shown with red crosses in all panels. **(A)** Gaussian kernel density (Equation 1, h=150bp) is computed at each read of the ChIP lane in the genomic region shown. **(B)** The derived ChIR is shown in black dots at each ChIP read.



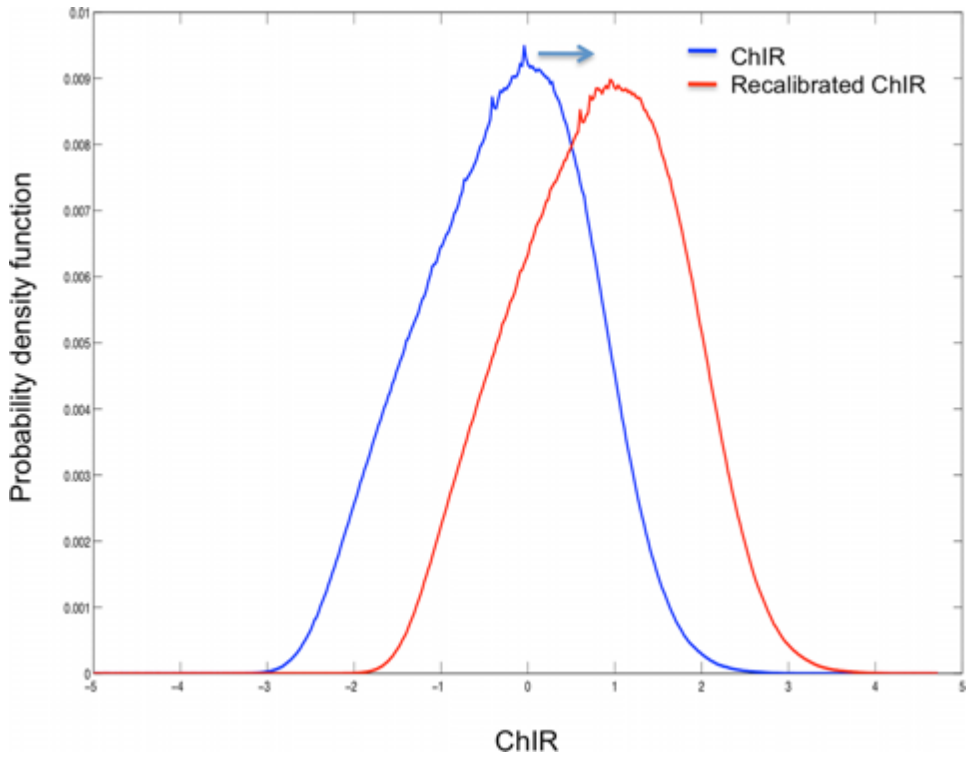
Supplementary Figure 5. Illustration of Gaussian kernel density and derived ChIR in genomic region chr7:52,968,000-52,981,000 of the MYO.SIN3B.GM ChIP-Seq dataset. The genomic positions of individual ChIP reads are shown with red crosses in all panels. **(A)** Gaussian kernel density (Equation 1, *h*=150bp) is computed at each read of the ChIP lane in the genomic region shown. **(B)** The derived ChIR is shown in black dots at each ChIP read.



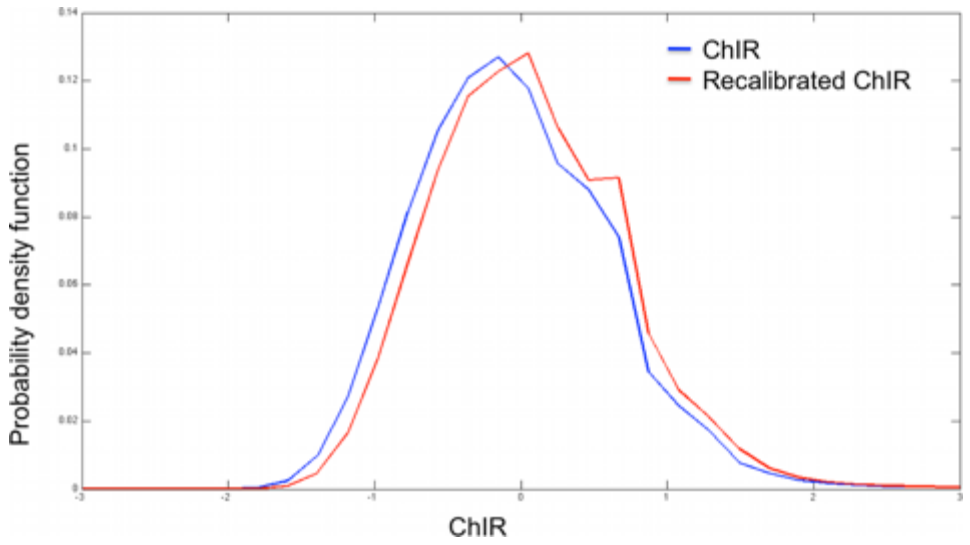
Supplementary Figure 6. Illustration of Qeseq's scheme in the same genomic region chr19: 54,605,999-54,617,335 of the MYO.H3K27me3.GM ChIP-Seq dataset. Qeseq's modules to identify a potential event are shown on chromosome 7. The horizontal axis shows genomic coordinates. **(A)** The green channel displays ChIP signal on the positive as well as negative strands. The individual green bars represent ChIP-Seq reads mapped to the given genomic location. Green bars sitting on top of each other indicate an accumulation of mapped reads in the given genomic region. **(B)** The red channel represents control (total input) reads on the positive as well as negative strands. The individual red bars represent control ChIP-Seq reads mapped to the given genomic location. **(C)** The blue channel displays ChIR on both strands of the ChIP lane. ChIR is calculated for each read in the ChIP lane according to Equation 2. Blue bars above the grey line are positive and indicate genomic regions where the ChIP lane has more reads than the control lane. Blue bars below the grey line are negative and mark genomic regions where the control lane has more reads than the ChIP lane. **(D)** The black channel displays in $\log_{10}$ scale the p-values associated with each ChIP-Seq read of the ChIP lane. Large negative bars indicate significant p-values, while black bars close to 0 are not significant. **(E)** Qeseq's cluster detection module identifies the boundaries of a potential event by joining neighbors with p-values <0.05 and/or joining neighbors less than the experimental fragment length (h) away. The black line represents reads that are merged together according to the cluster detection module. **(F)** Recalibrated ChIR of the potential event is displayed in the purple channel. The purple bars represent the recalibrated ChIR associated with each read in the ChIP lane. **(G)** Final location of event is marked by the brown rectangle. This figure has been generated using the IGB genome browser (http://bioviz.org/igb/).



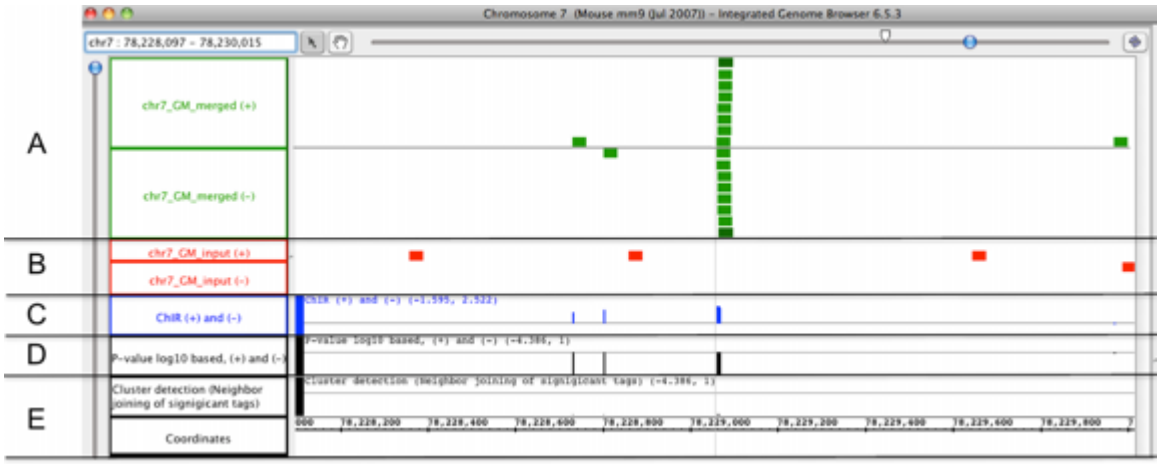
Supplementary Figure 7. Illustration of Qeseq's scheme in the same genomic region chr7:52,968,000-52,981,000 of the MYO.SIN3B.GM ChIP-Seq dataset. Qeseq's modules to identify a potential event are shown on chromosome 7. The horizontal axis shows genomic coordinates. **(A)** The green channel displays ChIP signal on the positive as well as negative strands. The individual green bars represent ChIP-Seq reads mapped to the given genomic location. Green bars sitting on top of each other indicate an accumulation of mapped reads in the given genomic region. **(B)** The red channel represents control (total input) reads on the positive as well as negative strands. The individual red bars represent control ChIP-Seq reads mapped to the given genomic location. **(C)** The blue channel displays ChIR on both strands of the ChIP lane. ChIR is calculated for each read in the ChIP lane according to Equation 2. Blue bars above the grey line are positive and indicate genomic regions where the ChIP lane has more reads than the control lane. Blue bars below the grey line are negative and mark genomic regions where the control lane has more reads than the ChIP lane. **(D)** The black channel displays in $\log_{10}$ scale the p-values associated with each ChIP-Seq read of the ChIP lane. Large negative bars indicate significant p-values, while black bars close to 0 are not significant. **(E)** Qeseq's cluster detection module identifies the boundaries of a potential event by joining neighbors with p-values <0.05 and/or joining neighbors less than the experimental fragment length (*h*) away. The black line represents reads that are merged together according to the cluster detection module. **(F)** Recalibrated ChIR of the potential event is displayed in the purple channel. The purple bars represent the recalibrated ChIR associated with each read in the ChIP lane. **(G)** Final location of event is marked by the brown rectangle. This figure has been generated using the IGB genome browser (http://bioviz.org/igb/).



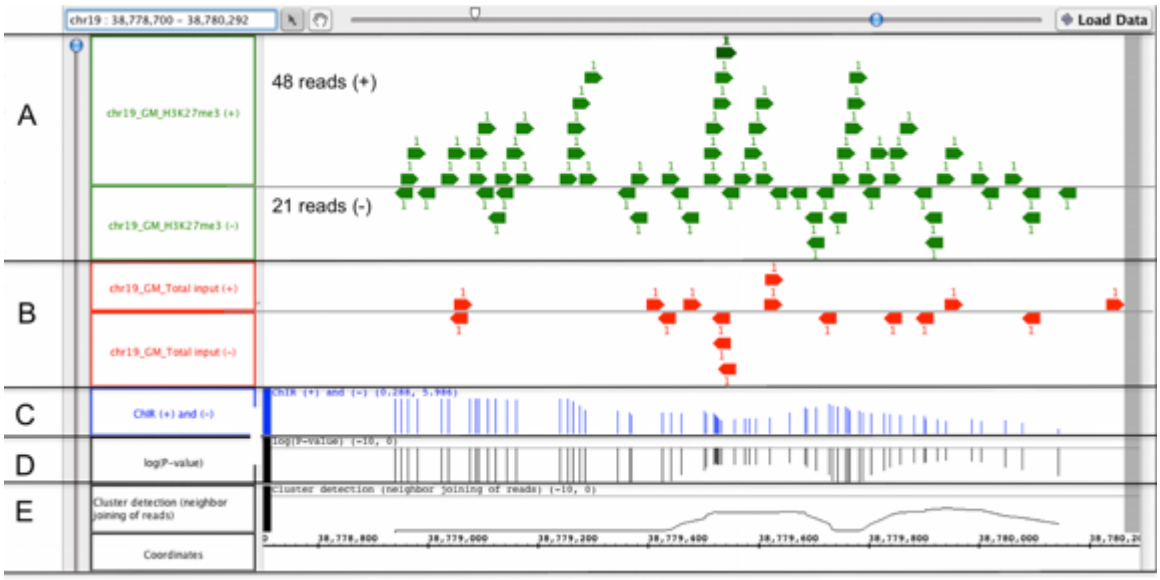
Supplementary Figure 8. Recalibration using the MYO.H3K27me3.GM ChIP-Seq dataset. Two ChIR distributions are shown; ChIR distribution before recalibration is represented by the blue color, while ChIR distribution after recalibration is displayed in red.



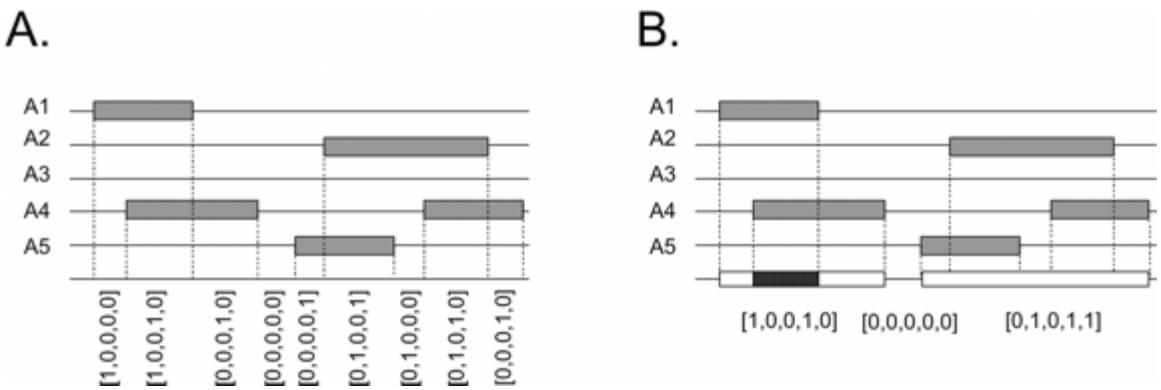
Supplementary Figure 9. Recalibration using the MYO.SIN3B.GM ChIP-Seq dataset. Two ChIR distributions are shown; ChIR distribution before recalibration is represented by the blue color, while ChIR distribution after recalibration is displayed in red. The recalibrated ChIR distribution is mean-centered at 0.



Supplementary Figure 10. Filtering of PCR artifacts in genomic region chr7:78,228,097-78,230,015 of the MYO.SIN3B.GM ChIP-Seq dataset. Qeseq's modules to identify a potential event are shown on chromosome 7. The horizontal axis shows genomic coordinates. **(A)** The green channel displays ChIP signal on the positive as well as negative strands. The individual green bars represent ChIP-Seq reads mapped to the given genomic location. Green bars sitting on top of each other indicate an accumulation of mapped reads in the given genomic region. **(B)** The red channel represents control (total input) reads on the positive as well as negative strands. The individual red bars represent control ChIP-Seq reads mapped to the given genomic location. **(C)** The blue channel displays ChIR on both strands of the ChIP lane. ChIR is calculated for each read in the ChIP lane according to Equation 2. Blue bars above the grey line are positive and indicate genomic regions where the ChIP lane has more reads than the control lane. Blue bars below the grey line are negative and mark genomic regions where the control lane has more reads than the ChIP lane. **(D)** The black channel displays in $\log_{10}$ scale the p-values associated with each ChIP-Seq read of the ChIP lane. Large negative bars indicate significant p-values, while black bars close to 0 are not significant. **(E)** Qeseq's cluster detection module identifies the boundaries of a potential event by joining neighbors with p-values <0.05 and/or joining neighbors less than the experimental fragment length (*h*) away. No final event is identified; the 'tower' of ChIP lane reads is identified as a potential PCR artifact. This figure has been generated using the IGB genome browser (http://bioviz.org/igb/).



Supplementary Figure 11. Filtering of unbalanced events in genomic region chr19: 54,605,999-54,617,335 of the MYO.H3K27me3.GM ChIP-Seq dataset. Qeseq's modules to identify a potential event are shown on chromosome 7. The region of interest is marked by a black vertical rectangle. The horizontal axis represents genomic coordinates. **(A)** The green channel displays ChIP signal on the positive as well as negative strands. The individual green bars represent ChIP-Seq reads mapped to the given genomic location. Green bars sitting on top of each other indicate an accumulation of mapped reads in the given genomic region. **(B)** The red channel represents control (total input) reads on the positive as well as negative strands. The individual red bars represent control ChIP-Seq reads mapped to the given genomic location. **(C)** The blue channel displays ChIR on both strands of the ChIP lane. ChIR is calculated for each read in the ChIP lane according to Equation 2. Blue bars above the grey line are positive and indicate genomic regions where the ChIP lane has more reads than the control lane. Blue bars below the grey line are negative and mark genomic regions where the control lane has more reads than the ChIP lane. **(D)** The black channel displays in $\log_{10}$ scale the p-values associated with each ChIP-Seq read of the ChIP lane. Large negative bars indicate significant p-values, while black bars close to 0 are not significant. **(E)** Qeseq's cluster detection module identifies the boundaries of a potential event by joining neighbors with p-values <0.05 and/or joining neighbors less than the experimental fragment length (h) away. No final event is identified as the negative strand of the ChIP lane contains 7 reads, while the positive strand only 2 reads. This figure has been generated using the IGB genome browser (http://bioviz.org/igb/).



Supplementary Figure 12. Generation of possible sites for validations using predictions from different algorithms. (A) Five algorithms are displayed as A1 to A5 on the Y-axis. The detected binding sites of each algorithm are shown as grey boxes. The genome is partitioned into intervals based on the predicted sites of these five algorithms, denoted by vertical lines. For each interval a fingerprint containing which algorithm predicts an event are computed. If an algorithm predicts an event in the particular interval, the fingerprint contains 1, otherwise 0. Algorithm 3 (A3) in this example does not predict any sites. **(B)** Clean-sites generated by the regions obtained by merging overlapping predictions. The intersection of all predictions in a region is shown in dark grey when present.

References

4. Laajala, T.D., Raghav, S., Tuomela, S., Lahesmaa, R., Aittokallio, T. and Elo, L.L. (2009) A practical comparison of methods for detecting transcription factor binding sites in ChIP-seq experiments. *Bmc Genomics*, **10**, 618.
5. Wilbanks, E.G. and Facciotti, M.T. (2010) Evaluation of Algorithm Performance in ChIP-Seq Peak Detection. *Plos One*, **5**, e11471.
16. Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M. and Gerstein, M.B. (2009) PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nature Biotechnology*, **27**, 66-75.
34. van Oevelen, C., Bowman, C., Pellegrino, J., Asp, P., Cheng, J., Parisi, F., Micsinai, M., Kluger, Y., Chu, A., Blais, A. et al. (2010) The mammalian Sin3 proteins are required for muscle development and sarcomere specification. *Mol Cell Biol*, **30**, 5686-5697.