# Application of NanoString nCounter technology to the Validation of ChIP-Seq datasets in the ENCODE project

Charles B. Epstein[1], Alon Goren[1,3,4], Melissa Gymrek, Jason Ernst[1,2], Noam Shoresh[1], Xiaolan Zhang[1], Robbyn Issner[1], Michael Coyne[1], Ido Amit[1], Aviv Regev[1], Manolis Kellis[1,2], Bradley E. Bernstein[1,3,4]*

1 Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

2 MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, Massachusetts, USA.

3 Howard Hughes Medical Institute, Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA

4. Center for Systems Biology and Center for Cancer Research, Massachusetts General Hospital, Boston, Massachusetts, USA

* Correspondence should be addressed to B.E.B. (Bernstein.Bradley@mgh.harvard.edu).

# Abstract

Highly multiplexed measurement of the abundances of specific nucleic acid sequences is a main-stay of the field of genomics. Here we extend the NanoString nCounter technology to the evaluation of fragment abundances following chromatin immunoprecipitation. We describe the method used to design a custom NanoString codeset for the purposes of measuring DNA fragment abundances derived from chromatin immunoprecipitation performed with a diverse collection of antibodies against histone modifications, as well as RNA Pol II and the factor CTCF. We then employ the custom NanoString array to measure fragment abundances from a total of 13 ChIPs and 3 libraries derived from ChIPs. We present graphically the correlation between the quantification of fragment abundances by the two methods. Correlations ranged from a high of 0.92 (H3K9ac) to a low of 0.09 (whole cell extract). Eleven of the 16 ChIP-seq – NanoString combinations tested had a correlation of at least 0.70. Overall, this report demonstrates that the inference of fragment relative abundances from ChIP-Seq can be confirmed by an independent method, thus validating our ChIP-Seq results in the ENCODE project.

# Introduction

The study of gene structure and regulation, and in particular the field of epigenomics, relies upon the mapping of the genomic locations of specific histone modifications and chromatin associated proteins. To accomplish this, chromatin is immunoprecipitated with appropriate antibodies, and the recovered DNA fragments are sequenced using Illumina short read sequencing. The short reads are aligned to a genome scaffold to discover sites of significant enrichment over background. Although this approach offers many advantages, including nucleotide scale resolution and the opportunity to do experiments without the prior design of an array of detectors, it suffers from the disadvantages of relatively high cost and low throughput.

In order to overcome these disadvantages, we adapted NanoString's nCounter technology to the measurement of locus specific enrichments following chromatin immunoprecipitation. In order to accomplish this, we had to solve two problems. First, we had to develop an experimental strategy for analyte labeling consistent with nCounter technology. Since nCounter was originally developed to be used with measurement of RNA species abundance, we adapted the manufacturer's supplied labeling technology to work with DNA fragments. Secondly, we had to design an informative collection of probes to use with the DNA fragment pools derived by immunoprecipitation of chromatin with antibodies targeting a range of selected histone modifications. Since the particular histone modifications of interest to us are associated with canonical features of gene structure, we created a probe array that sampled such features, including enhancers, promoters, transcription start sites, and transcribed or potentially imprinted regions.

This report will present in detail how the array was designed, explore its baseline behavior in terms of reproducibility, background, cross hybridization artifacts, and dynamic range, and show the comparison between results obtained using the routine methods of ChIP-Seq and the novel method of ChIP-String. Overall, we found excellent correlation between genome scale measurements made using ChIP-Seq and NanoString, thus validating our ChIP-Seq data in the Encode project.

# Methods

### Nanostring code set design

Two independent 10 state Chrom-HMM analyses (Ernst and Kellis, Nature Biotech, 2010; Ernst et al, Nature, 2011) were conducted using ChIP-Seq data from two cell types (K562 and H1 ESCs), using all ChIP-seq tracks available at the time. The ChIP-Seq tracks employed are summarized in Figure 1. For each state in each cell type (20 states, total), fifty 400mers were sampled to represent genomic regions associated with the corresponding state. The sampling was done based on the posterior distribution for the state, restricted to regions where all 8-25bp intervals in the 400bp interval were uniquely alignable to the HG18 reference genome. The entire set of one thousand 400mers was supplied to Nanostring Technologies for probe design, i.e. the process of identification of an optimal 100mer (if any exist) within the 400mer that could potentially be used as a probe in a Nanostring codeset, satisfying the criteria of uniqueness in the genome and mutual independence from the other probes in the design set. This led to the design of 903 candidate probes, with zero or one coming from each of the original 1000 regions.

We took the set of 903 candidate probes, and for each available ChIP-Seq dataset in 8 cell types (some 90 datasets in all), we tabulated the ChIP-Seq scaled intensity over all 90 datasets. The 8 cell types included the 2 used in the original HMM, plus six other cell types used in the ENCODE project (HepG2, HMEC, HUVEC, NHEK, NHLF, GM12878). We then analyzed the 90 x 903 matrix to identify an optimally sensitive and diverse set of 490 probes, using the following method.

To overcome the challenge posed by the fact that many of the 903 regions showed little or no enrichment in our prior ChIP-Seq results (hence were expected to yield only background readings in a Nanostring assay, regardless of the antibody employed in the ChIP), we ranked the regions and selected the 490 regions expected to give the strongest detection over the full range of epitopes employed in assay design. The ranking was performed as follows: For each of the 9 ChIP-Seq epitopes consistently represented in our 90 datasets, we computed the average ChIP-Seq signal over the 8 cell types. Using the average performance over the 8 cells types, we ranked each probe with a number from 1(high signal) to 903 (low signal) for each epitope. We then computed the minimum value of the ranks for each probe over the 9 epitopes. We then identified within each HMM class the 24 or 25 probes having the lowest minimum ranks, indicating that these were the best probes to detect the strongest possible signal from one or more of the ChIP-Seq antibodies used in prior data. Note, to preserve the diversity of genomic regions represented on the array, we retained nearly equal numbers (24 or 25 members) of each of the 20 HMM states identified in the initial step. This was necessary because some HMM classes are intrinsically “dim”, meaning they correspond to regions that were not particularly enriched in any known ChIP-Seq datasets at the time the analysis was conducted.

An additional 7 probes were designed to encompass regions representing Zinc finger genes and known sites of H3K9me3 enrichment, such as the KCNQ1 locus. Finalization of probe selection was accomplished in collaboration with Nanostring technologies based on further constraints on probe performance discovered during synthesis and testing. The final resulting Nanostring array had 487 probes (in the first synthesis batch) and 489 probes in subsequent batches, from the initial 497 designs. One probe (B24) was omitted from all analyses when it was found to correspond to a repetitive DNA element, hence the binning analysis described below employed 486 probes. The sequence coordinates of the full set of 487 probes are found in the accompanying Excel spread sheet.

Figure 2 illustrates the average ChIP-Seq intensity over all 50 originally designed probes associated with a particular HMM class, for each epitope, from prior ChIP-Seq data available during probe design and selection. As anticipated from the figure, probes derived from certain HMM classes, such as B and L (from H1 cells and K562 cells, respectively) gave rise to bright signals in H3K4me3, H3K4me2, and H3K9ac ChIPs, whereas probes derived from other HMM classes, such as H and R, at most weakly detect signals in nearly all assays.

### Labeling sample for measurement

ChIP DNA (5 to 100 ng) was brought to a volume of 5ul and thermally denatured at 95C for 5', then immediately transferred to ice. DNA was supplemented with 10 ul of Reporter CodeSet and 10 ul of Hybridization Buffer (supplied by Nanostring Technologies). When ready, 5 ul of Capture ProbeSet was added, and the sample was immediately placed at 65C overnight (> 12 hours). Following overnight incubation, samples are automatically processed in a Nanostring Prep Station following the manufacturer's protocol. The processed sample is then imaged in a Nanostring Imaging station.

### Analysis of Nanostring assay results

Data are linearly rescaled (normalized) using the supplied exogenous positive controls, employing the convention that the 6 positive controls sum to a value of 60,000 after rescaling. Typical scale factors are in the range from 0.5 to 5, and the median scale factor was 1.74 over the first 200 assays run in our lab. To explore graphically the correspondence between NanoString results and ChIP-Seq results, the following procedure was followed: For each ChIP-Seq data set, numerical values were derived for each of 486 genomic regions in the NanoString codeset, and divided by the genome wide average value of a region of equivalent size from the same dataset, to approximate the fold enrichment above background of each region. The 486 values were then placed in "bins" based on the following bin – range boundaries: Under 2 fold enrichment, 2 – 5 fold enrichment, 5 – 10 fold enrichment, 10 – 20 fold enrichment, 20 – 40 fold enrichment, and > 40 fold enrichment. The NanoString assay values for the probes corresponding to each bin were statistically summarized as 5th, 25th, median, 75th, and 95th percentiles, and the statistical range of NanoString assay results were plotted against the bin identifier. The Y axis was re-scaled so that the median value of the bin corresponding to < 2 fold enrichment = 100 arbitrary units. The resulting graphs are attached.

# Results

## 1. Distribution and reproducibility of background values

As a preliminary experiment, we wished to measure the intrinsic signal associated with each probe in the absence of input DNA. We performed a labeling assay with no input DNA, and measured (with four fold replication) the distribution of probe counts (intensities). We found that the median normalized assay value was 15.7, which is appropriately low, considering that a true signal in the assay can be in the range of several thousands. However, we also found that the background values for each probe reproducibly fell in some range, and that two probes at the high end of the range (L43 and B40) were high enough to be of some concern, giving mean assay values of 119 and 95, respectively (Figure 3). The very weak, but reproducible, signals in the absence of true signal must be taken into account in large scale studies that employ this platform, insofar as spurious inferences of similarity could be drawn if multiple assays were run using samples having essentially no signal. We also found a typical inverse relation between assay value and CV, as shown in Figure 4. Thus, while for unknown reasons, some probes have a reproducibly high background, this background is still trivial compared to the values obtained in a true assay.

## 2. Reproducibility of the assay

We used the ChIP-String assay to make replicate measurements of 5 ng of ChIP DNA, derived using K562 cells and H3K4me3 antibody. We found that ChIP-String assay values were highly reproducible between assays, in spite of the very low amount of input DNA employed, giving a correlation (r squared) of 0.94; correlation plot is shown in Figure 5. The average assay values were 112 and 232, and the maximal assay values were 2550 and 6612. One probe (B24) was found to give a very high value in all assays of human DNA and is omitted from this and all further calculations. This probe was later found to interact with a repetitive DNA element, in spite of the otherwise successful effort to exclude such probes during the design of the probe array.

## 3. Correlation  of Nanostring results from assays of a ChIP and a library derived from that ChIP

We compared 5 ng of H3K4me3 ChIP DNA to 33 ng of library DNA derived from that ChIP, using simple correlation analysis (Figure 6). We found excellent agreement between the two, with an overall R squared of 0.89. The assay of the library gave much higher ChIP-String counts than the assay of the ChIP, consistent with the greater amount of input DNA. Thus, we have demonstrated, in a highly multiplexed assay, the overall agreement between fragment abundances in a ChIP, and in a library derived from that ChIP. We repeated the comparison of NanoString assays of ChIP and library using H3K9ac (Fig. 7), H3K27ac (Fig. 8), H3K27me3 (Fig. 9), and H3K36me3 (Fig. 10). As expected, the agreements are superior for the marks having higher signal to background ratios, but overall, the agreement between corresponding ChIP and library is extremely good. The agreement between ChIP and library, which is a foundation of the accuracy of the ChIP-Seq method, implies that one does not introduce significant bias when generating libraries from ChIPs. While heretofore this has been demonstrated using comparisons of ChIP-on-chip and ChIP-seq, for low resolution, and using qPCR of ChIPs and libraries, for high resolution, we are not aware of any prior high resolution demonstrations of the validity of ChIP-seq which approach the scale of the present report.

## 4. Validation of ENCODE project ChIP-Seq data by comparison of ChIP-Seq results with ChIP-String results

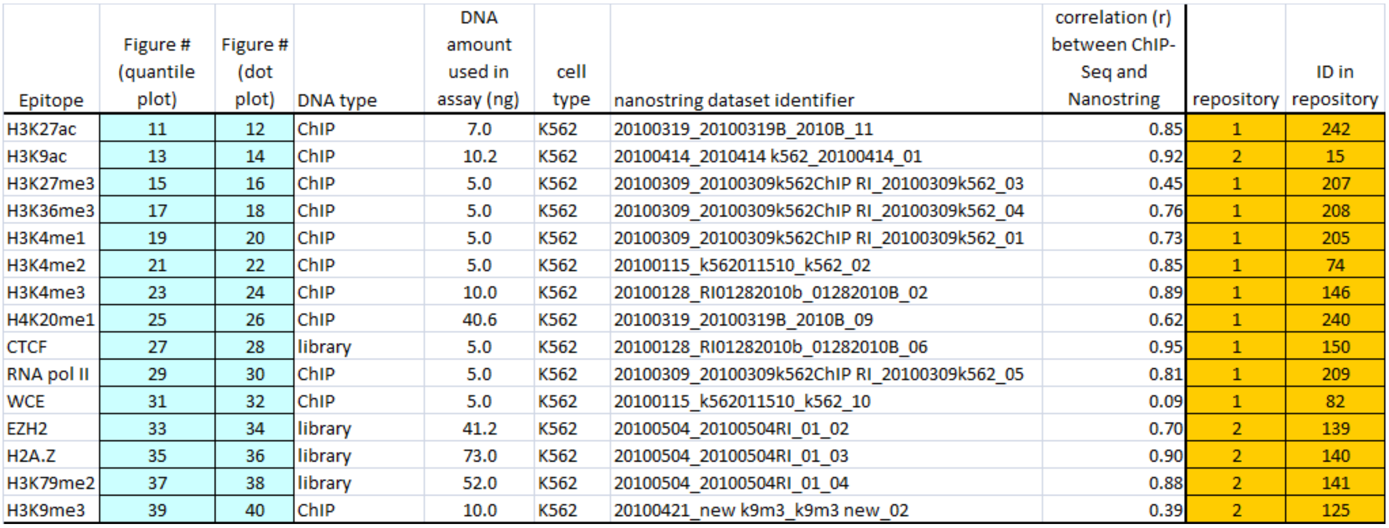
In order to **validate** the ChIP-Seq data generated by our group in the ENCODE project, we derived new ChIPs from an independent culture of K562 cells, and performed ChIP-String assays on each ChIP. We then compared the resulting ChIP-String data to the previously obtained ChIP-Seq data. We explored several methods for comparing ChIP-String and ChIP-Seq data; in each case, we compared a derived metric (explained below) to the ChIP-Seq integrated tag count from the genomic region corresponding to each Nanostring probe. The following derived metrics were evaluated:

(a) ratio of Nanostring measurements of ChIP to Nanostring measurements of input control;
(b) ratio of Nanostring measurements of ChIP to Nanostring measurements of input control, with a floor value of 20 substituted when necessary; and
(c) normalized Nanostring measurements of ChIP used without any adjustment based on input control.

We found that input controls and ChIPs were uncorrelated in Nanostring measurements, so methods (a) and (b) were noisy compared to method (c). All the comparison plots referenced below used method (c), which simply compares the normalized Nanostring measurement, probe by probe, to the integrated ChIP-Seq tag count in the region covered by the probe.

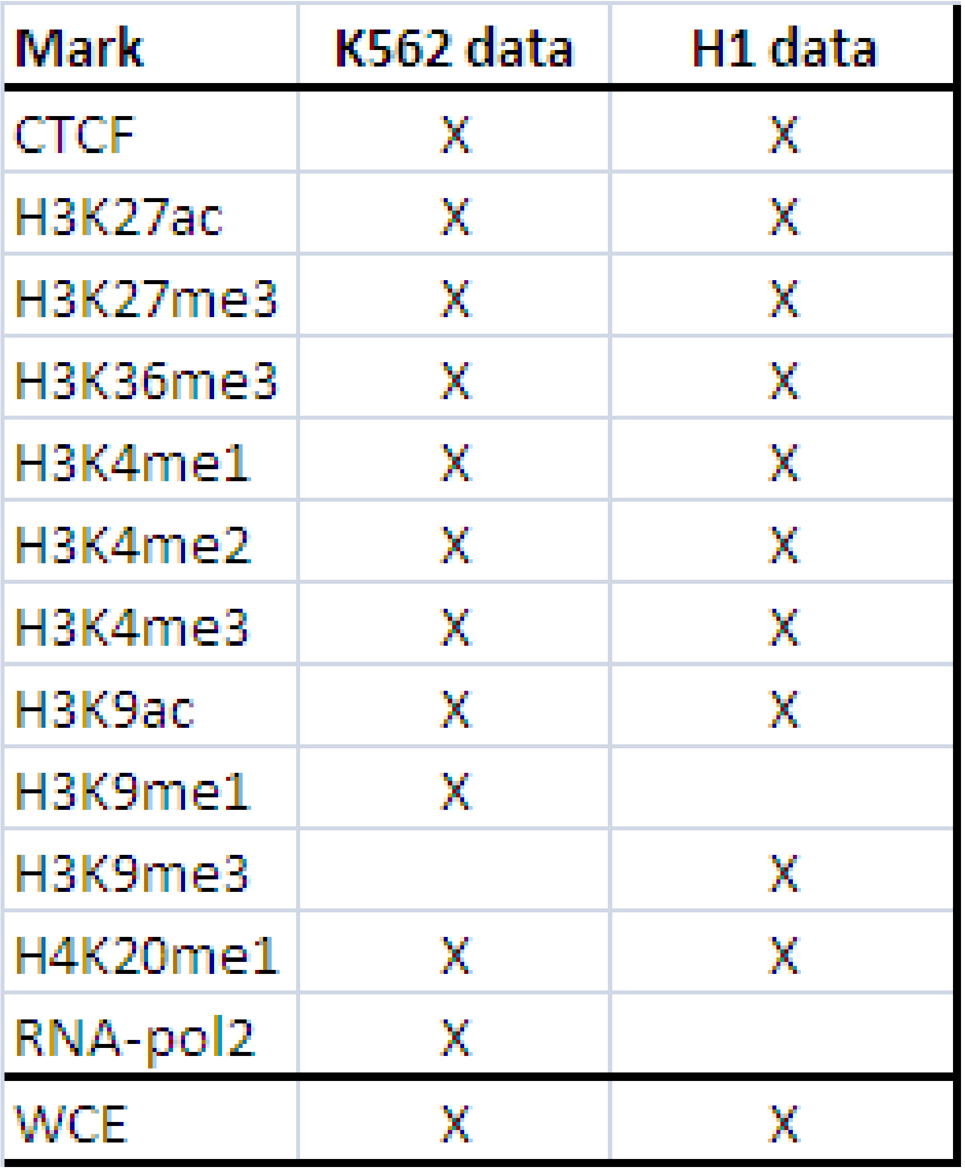
Comparisons were done using both dot plots (i.e. graphs with one points per probe), and using a statistical summary of the range of Nanostring assay values obtained on a quantile by quantile basis, where each quantile corresponds to a range of ChIP-Seq fold enrichments (signal over background). Results for epitopes are found in the attached figures, numbered as summarized in the first two columns of the following table. The table also shows what type of DNA was analyzed by NanoString (ChIP or Illumina sequencing library derived from ChIP), the amount of DNA input to the NanoString assay, the unique identifier for the NanoString assay (for purposes of establishing correspondence with the tabulated NanoString data supplied in the accompanying spread sheet), and the correlation (r) between ChIP-Seq and NanoString over the entire 486 region code set.

### Table 1: Index to Graphs and other statistics



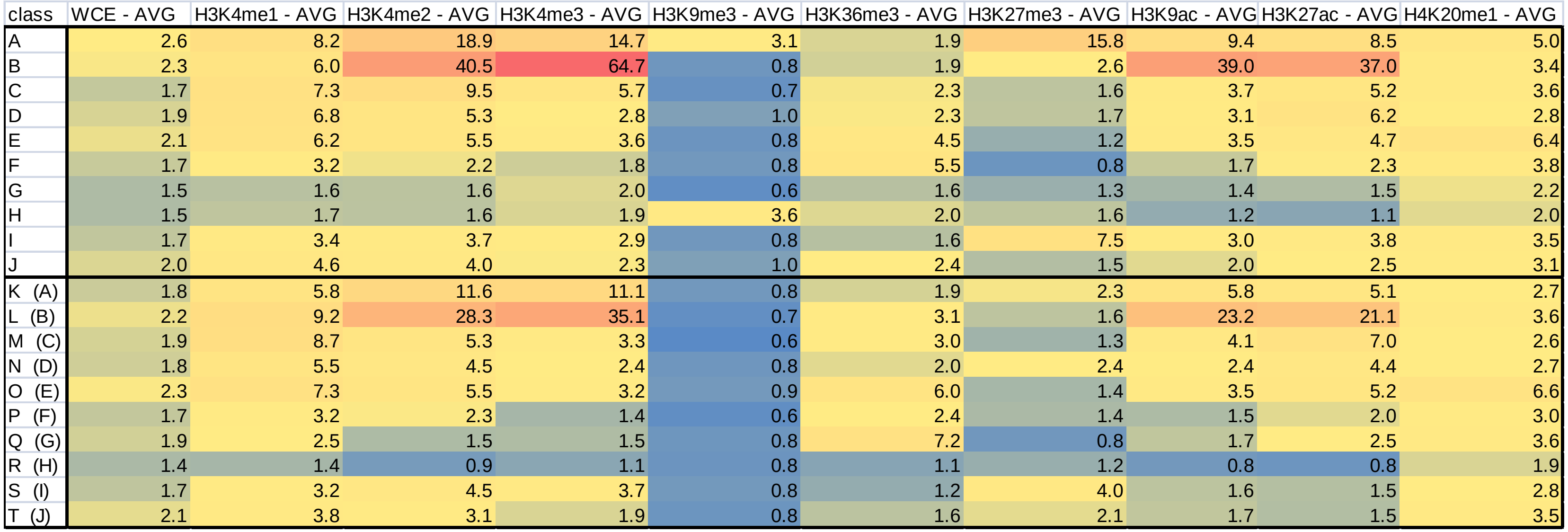
| Epitope | Figure # (quantile plot) | Figure # (dot plot) | DNA type | DNA amount used in assay (ng) | cell type | nanostring dataset identifier | correlation (r) between ChIP-Seq and Nanostring | repository | ID in repository |
|---|---|---|---|---|---|---|---|---|---|
| H3K27ac | 11 | 12 | ChIP | 7.0 | K562 | 20100319_20100319B_2010B_11 | 0.85 | 1 | 242 |
| H3K9ac | 13 | 14 | ChIP | 10.2 | K562 | 20100414_2010414 k562_20100414_01 | 0.92 | 2 | 15 |
| H3K27me3 | 15 | 16 | ChIP | 5.0 | K562 | 20100309_20100309k562ChIP RI_20100309k562_03 | 0.45 | 1 | 207 |
| H3K36me3 | 17 | 18 | ChIP | 5.0 | K562 | 20100309_20100309k562ChIP RI_20100309k562_04 | 0.76 | 1 | 208 |
| H3K4me1 | 19 | 20 | ChIP | 5.0 | K562 | 20100309_20100309k562ChIP RI_20100309k562_01 | 0.73 | 1 | 205 |
| H3K4me2 | 21 | 22 | ChIP | 5.0 | K562 | 20100115_k562011510_k562_02 | 0.85 | 1 | 74 |
| H3K4me3 | 23 | 24 | ChIP | 10.0 | K562 | 20100128_RI01282010b_01282010B_02 | 0.89 | 1 | 146 |
| H4K20me1 | 25 | 26 | ChIP | 40.6 | K562 | 20100319_20100319B_2010B_09 | 0.62 | 1 | 240 |
| CTCF | 27 | 28 | library | 5.0 | K562 | 20100128_RI01282010b_01282010B_06 | 0.95 | 1 | 150 |
| RNA pol II | 29 | 30 | ChIP | 5.0 | K562 | 20100309_20100309k562ChIP RI_20100309k562_05 | 0.81 | 1 | 209 |
| WCE | 31 | 32 | ChIP | 5.0 | K562 | 20100115_k562011510_k562_10 | 0.09 | 1 | 82 |
| EZH2 | 33 | 34 | library | 41.2 | K562 | 20100504_20100504RI_01_02 | 0.70 | 2 | 139 |
| H2A.Z | 35 | 36 | library | 73.0 | K562 | 20100504_20100504RI_01_03 | 0.90 | 2 | 140 |
| H3K79me2 | 37 | 38 | library | 52.0 | K562 | 20100504_20100504RI_01_04 | 0.88 | 2 | 141 |
| H3K9me3 | 39 | 40 | ChIP | 10.0 | K562 | 20100421_new k9m3_k9m3 new_02 | 0.39 | 2 | 125 |

# ChIP-Seq data sets employed in Chrom HMM



| Mark | K562 data | H1 data |
|---|---|---|
| CTCF | X | X |
| H3K27ac | X | X |
| H3K27me3 | X | X |
| H3K36me3 | X | X |
| H3K4me1 | X | X |
| H3K4me2 | X | X |
| H3K4me3 | X | X |
| H3K9ac | X | X |
| H3K9me1 | X | |
| H3K9me3 | | X |
| H4K20me1 | X | X |
| RNA-pol2 | X | |
| WCE | X | X |

# Average ChIP-Seq values from 10 chromatin marks computed from 8 cell types, corresponding to each of the 20 HMM states

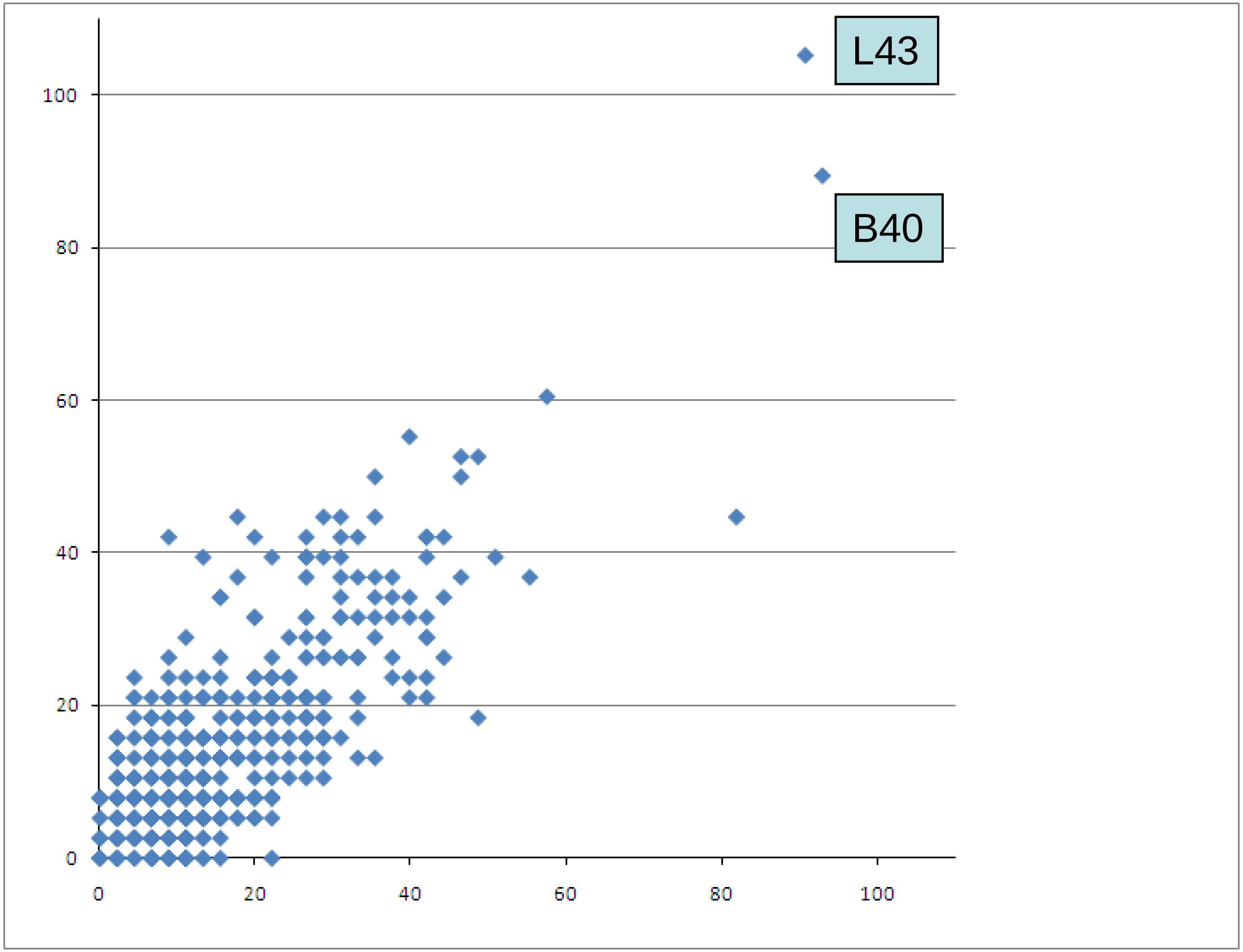


| class | WCE - AVG | H3K4me1 - AVG | H3K4me2 - AVG | H3K4me3 - AVG | H3K9me3 - AVG | H3K36me3 - AVG | H3K27me3 - AVG | H3K9ac - AVG | H3K27ac - AVG | H4K20me1 - AVG |
|---|---|---|---|---|---|---|---|---|---|---|
| A | 2.6 | 8.2 | 18.9 | 14.7 | 3.1 | 1.9 | 15.8 | 9.4 | 8.5 | 5.0 |
| B | 2.3 | 6.0 | 40.5 | 64.7 | 0.8 | 1.9 | 2.6 | 39.0 | 37.0 | 3.4 |
| C | 1.7 | 7.3 | 9.5 | 5.7 | 0.7 | 2.3 | 1.6 | 3.7 | 5.2 | 3.6 |
| D | 1.9 | 6.8 | 5.3 | 2.8 | 1.0 | 2.3 | 1.7 | 3.1 | 6.2 | 2.8 |
| E | 2.1 | 6.2 | 5.5 | 3.6 | 0.8 | 4.5 | 1.2 | 3.5 | 4.7 | 6.4 |
| F | 1.7 | 3.2 | 2.2 | 1.8 | 0.8 | 5.5 | 0.8 | 1.7 | 2.3 | 3.8 |
| G | 1.5 | 1.6 | 1.6 | 2.0 | 0.6 | 1.6 | 1.3 | 1.4 | 1.5 | 2.2 |
| H | 1.5 | 1.7 | 1.6 | 1.9 | 3.6 | 2.0 | 1.6 | 1.2 | 1.1 | 2.0 |
| I | 1.7 | 3.4 | 3.7 | 2.9 | 0.8 | 1.6 | 7.5 | 3.0 | 3.8 | 3.5 |
| J | 2.0 | 4.6 | 4.0 | 2.3 | 1.0 | 2.4 | 1.5 | 2.0 | 2.5 | 3.1 |
| K (A) | 1.8 | 5.8 | 11.6 | 11.1 | 0.8 | 1.9 | 2.3 | 5.8 | 5.1 | 2.7 |
| L (B) | 2.2 | 9.2 | 28.3 | 35.1 | 0.7 | 3.1 | 1.6 | 23.2 | 21.1 | 3.6 |
| M (C) | 1.9 | 8.7 | 5.3 | 3.3 | 0.6 | 3.0 | 1.3 | 4.1 | 7.0 | 2.6 |
| N (D) | 1.8 | 5.5 | 4.5 | 2.4 | 0.8 | 2.0 | 2.4 | 2.4 | 4.4 | 2.7 |
| O (E) | 2.3 | 7.3 | 5.5 | 3.2 | 0.9 | 6.0 | 1.4 | 3.5 | 5.2 | 6.6 |
| P (F) | 1.7 | 3.2 | 2.3 | 1.4 | 0.6 | 2.4 | 1.4 | 1.5 | 2.0 | 3.0 |
| Q (G) | 1.9 | 2.5 | 1.5 | 1.5 | 0.8 | 7.2 | 0.8 | 1.7 | 2.5 | 3.6 |
| R (H) | 1.4 | 1.4 | 0.9 | 1.1 | 0.8 | 1.1 | 1.2 | 0.8 | 0.8 | 1.9 |
| S (I) | 1.7 | 3.2 | 4.5 | 3.7 | 0.8 | 1.2 | 4.0 | 1.6 | 1.5 | 2.8 |
| T (J) | 2.1 | 3.8 | 3.1 | 1.9 | 0.8 | 1.6 | 2.1 | 1.7 | 1.5 | 3.5 |

printed 1/3/2012

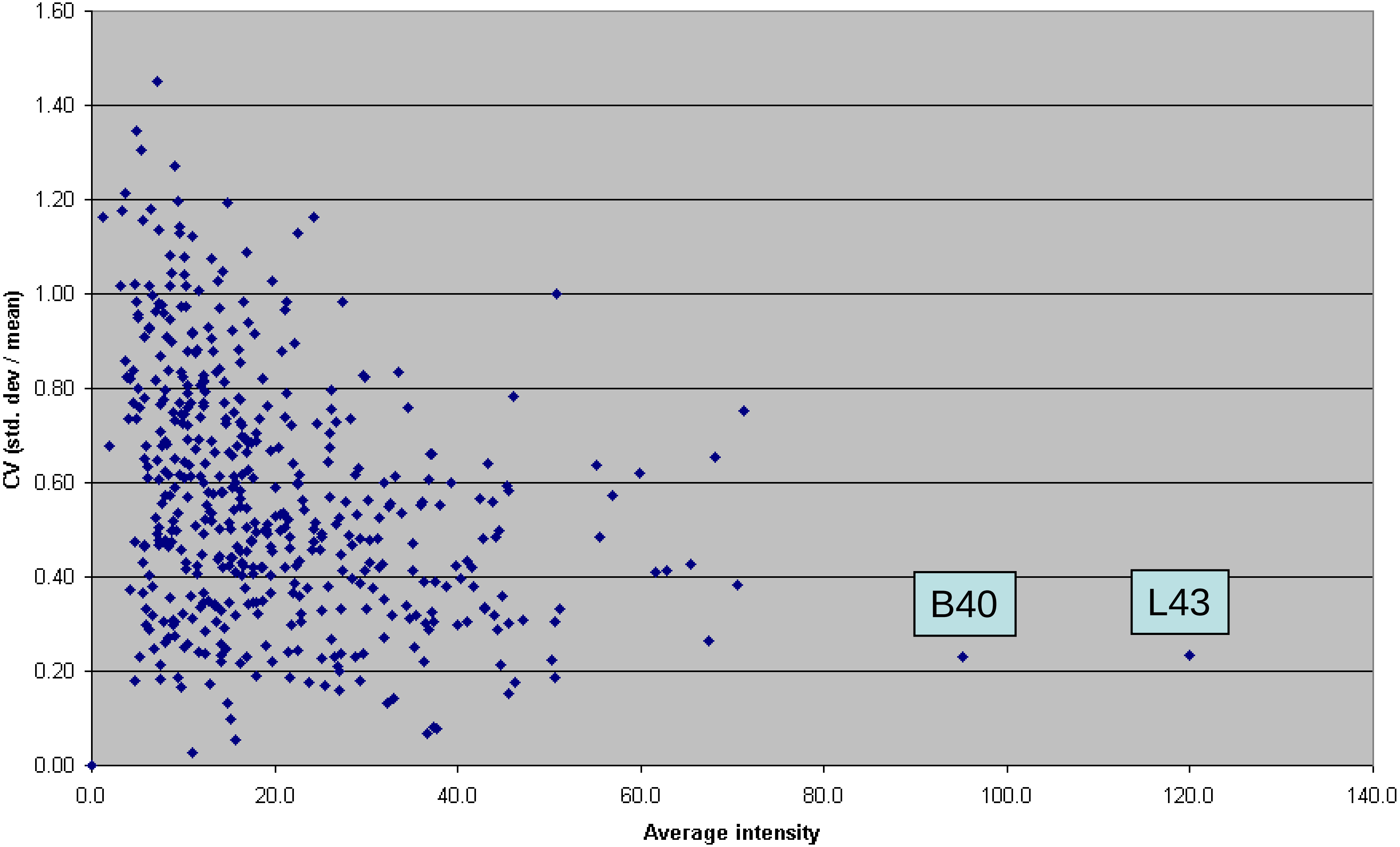
HMM classes A through J are based on the HMM inferred from H1 cell data, and HMM classes K through T are based on the HMM inferred from K562 cells. Blue corresponds to low values, yellow to intermediate values, and red corresponds to high values. HMM class R corresponds to regions that are low intensity in nearly every ChIP (horizontal blue band).

# Correlation of replicate assays with no input DNA



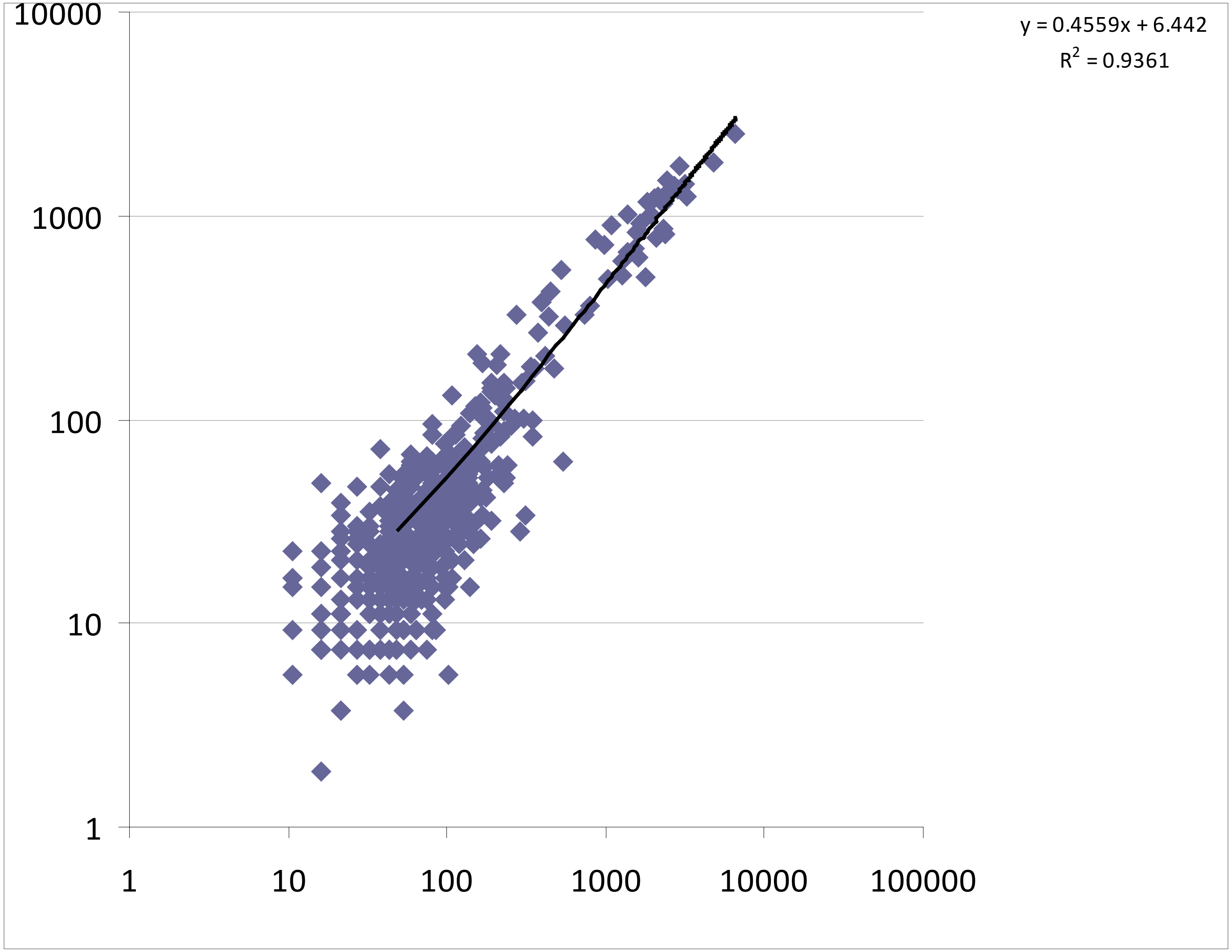
printed 1/3/2012

CV vs. intensity in replicate measurements of water



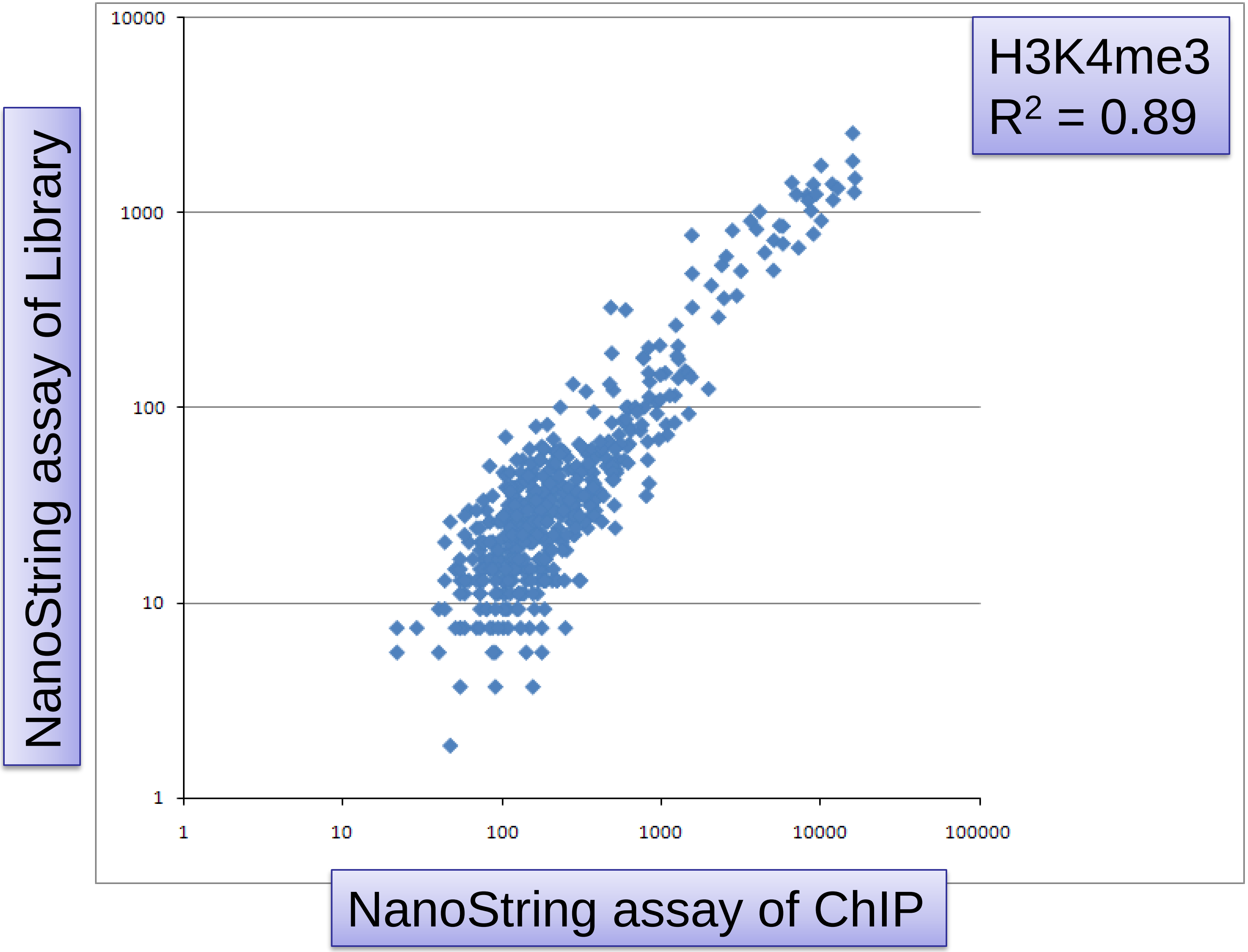
printed 1/3/2012

# Correlation of Replicate Assays of ChIP DNA



printed 1/3/2012

# Correlation of Assays of ChIP and Library DNA



H3K4me3

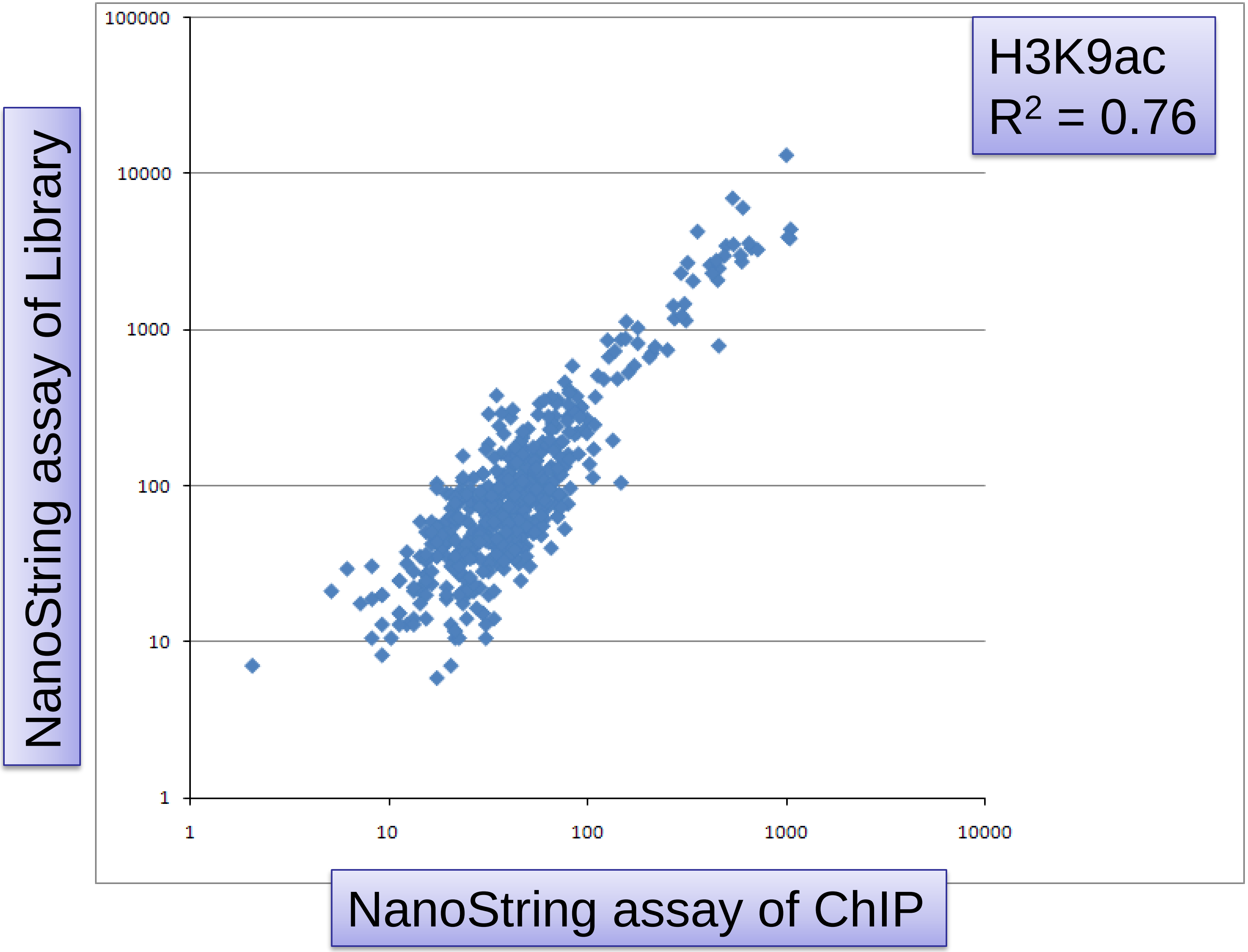
$R^2 = 0.89$

NanoString assay of Library

NanoString assay of ChIP

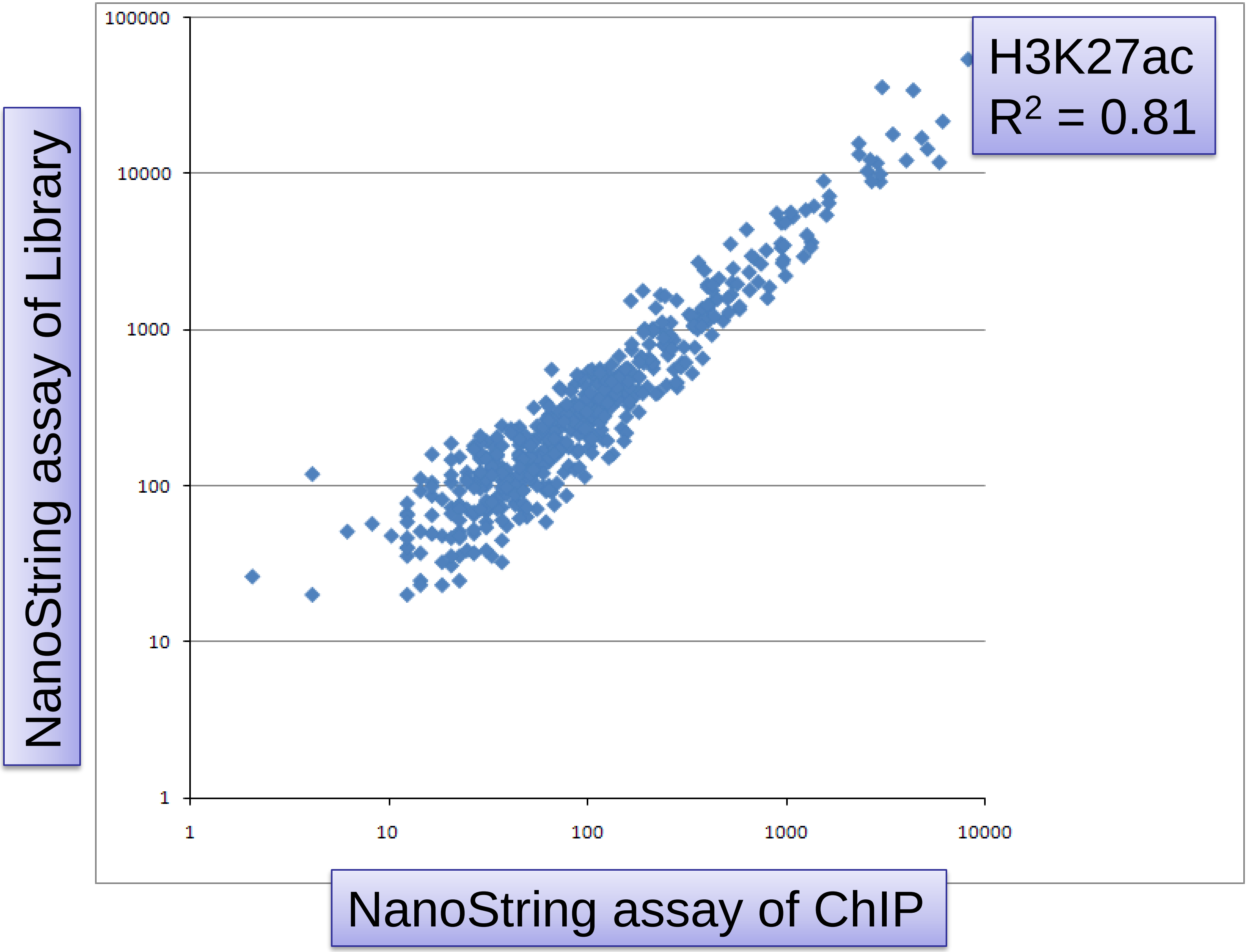
printed 1/3/2012

# Correlation of Assays of ChIP and Library DNA



printed 1/3/2012

# Correlation of Assays of ChIP and Library DNA

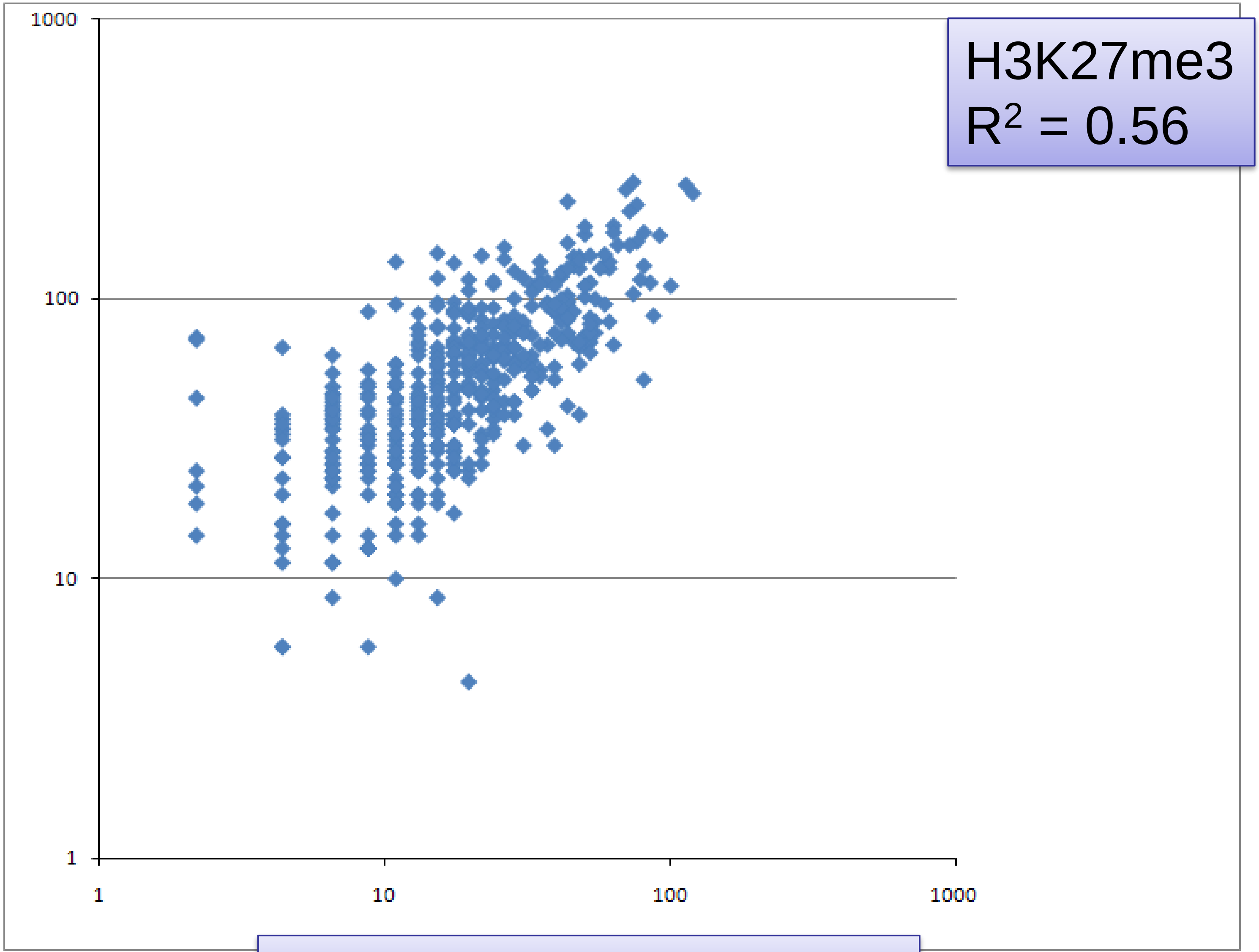


H3K27ac

$R^2 = 0.81$

printed 1/3/2012

# Correlation of Assays of ChIP and Library DNA



H3K27me3

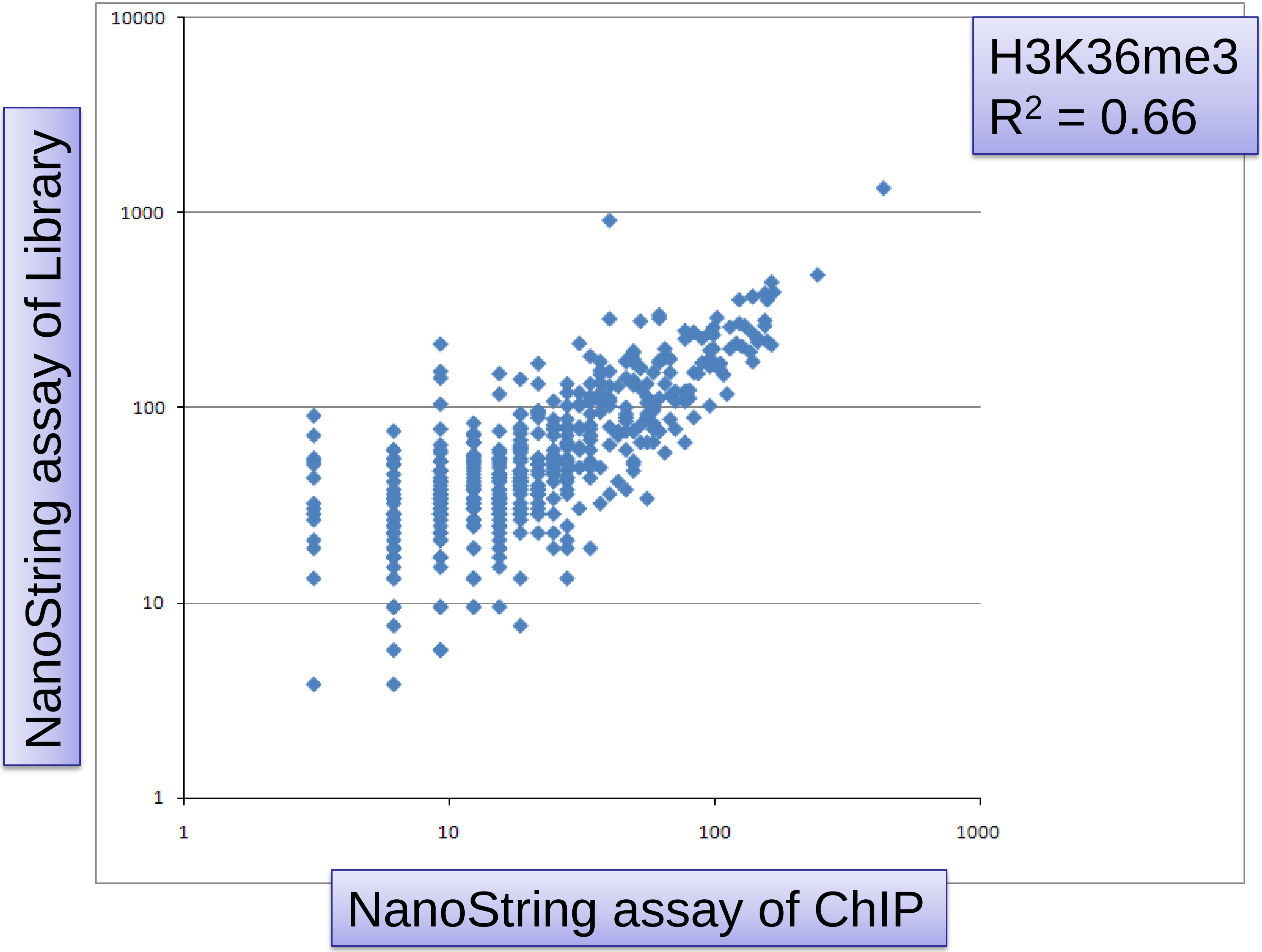
$R^2 = 0.56$

NanoString assay of Library

NanoString assay of ChIP

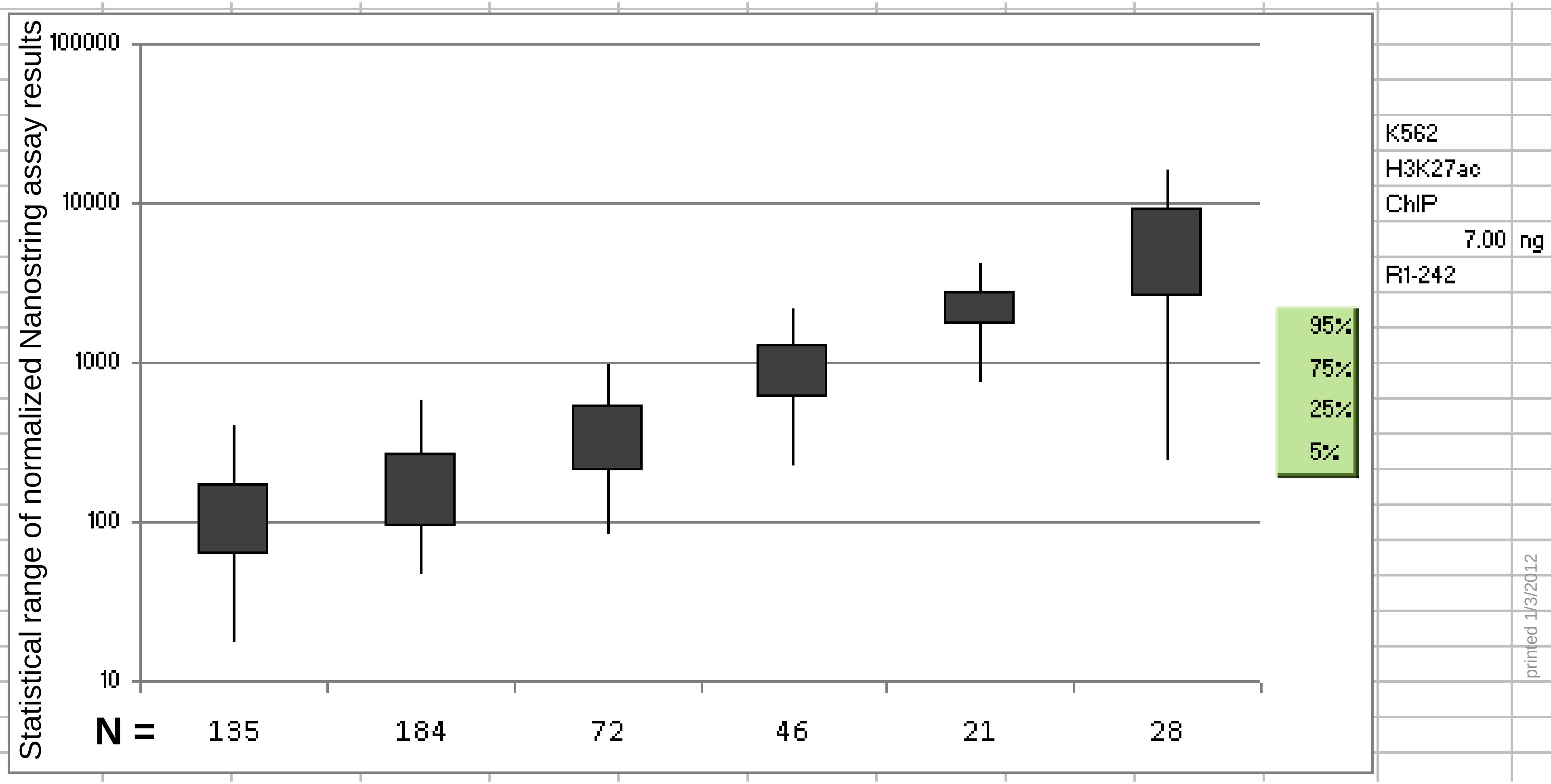
printed 1/3/2012

# Correlation of Assays of ChIP and Library DNA



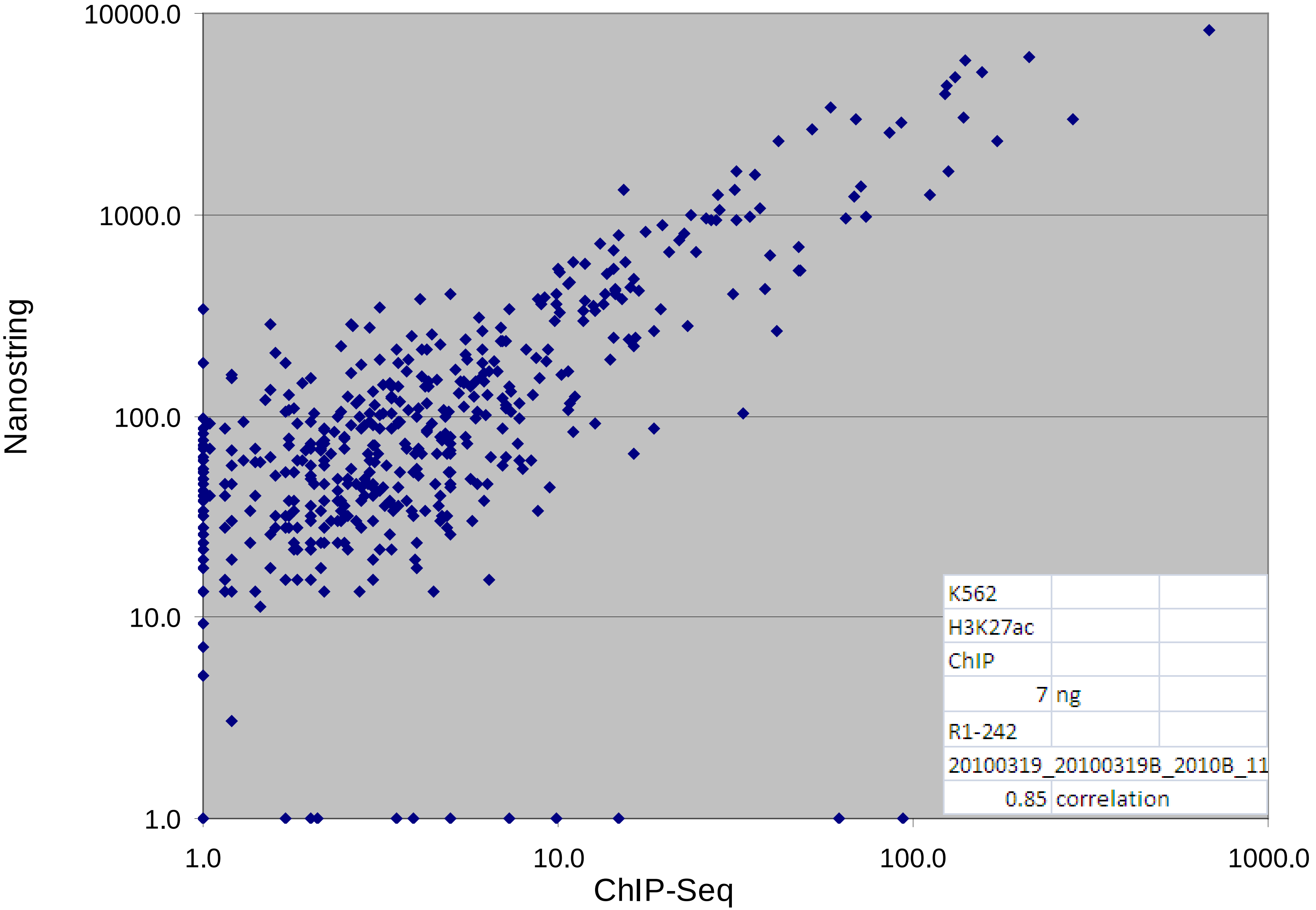
printed 1/3/2012

# K562 H3K27ac ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

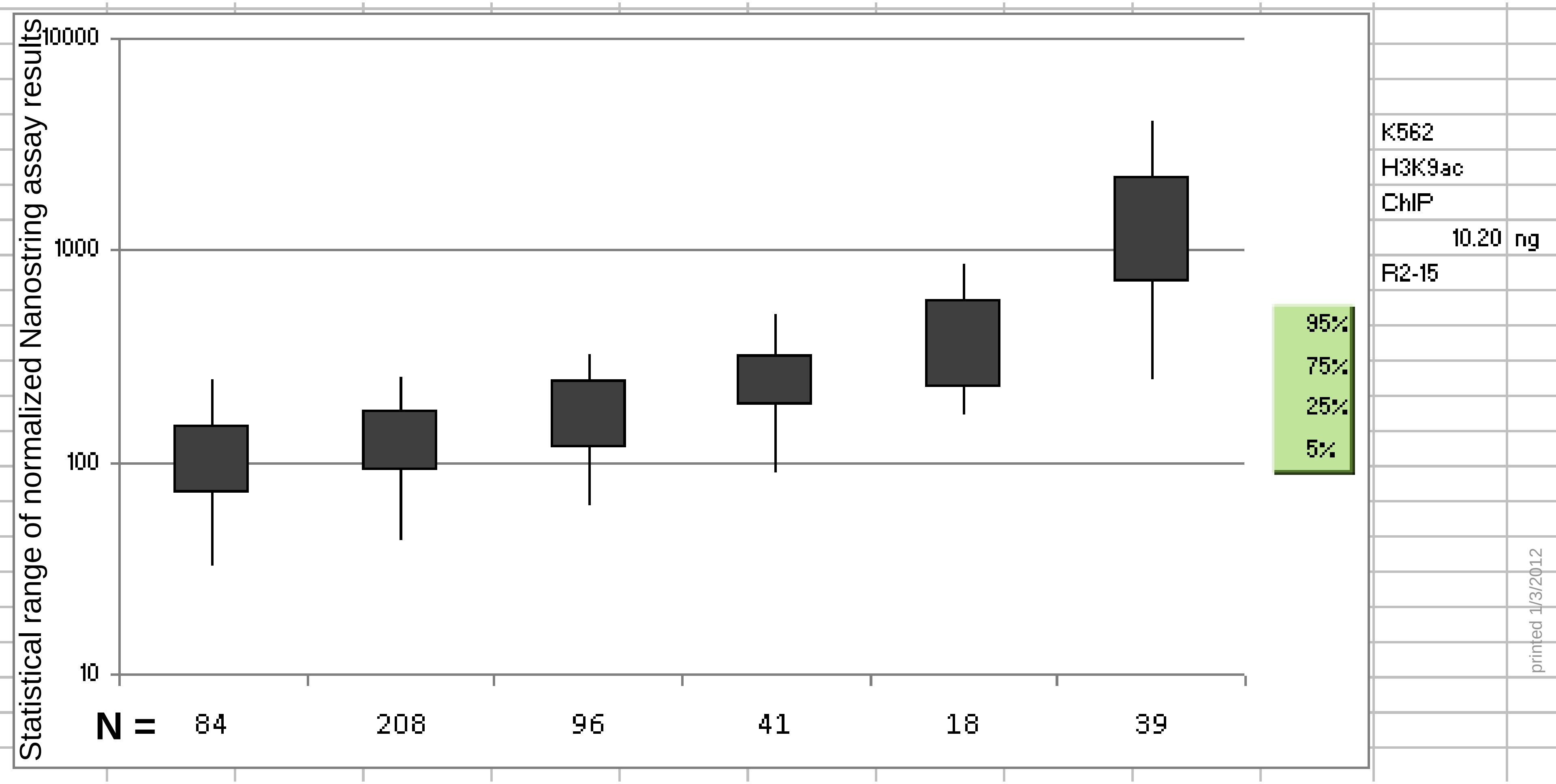
# K562 H3K27ac ChIP



| | | |
|---|---|---|
| K562 | | |
| H3K27ac | | |
| ChIP | | |
| 7 | ng | |
| R1-242 | | |
| 20100319_20100319B_2010B_11 | | |
| 0.85 | correlation | |

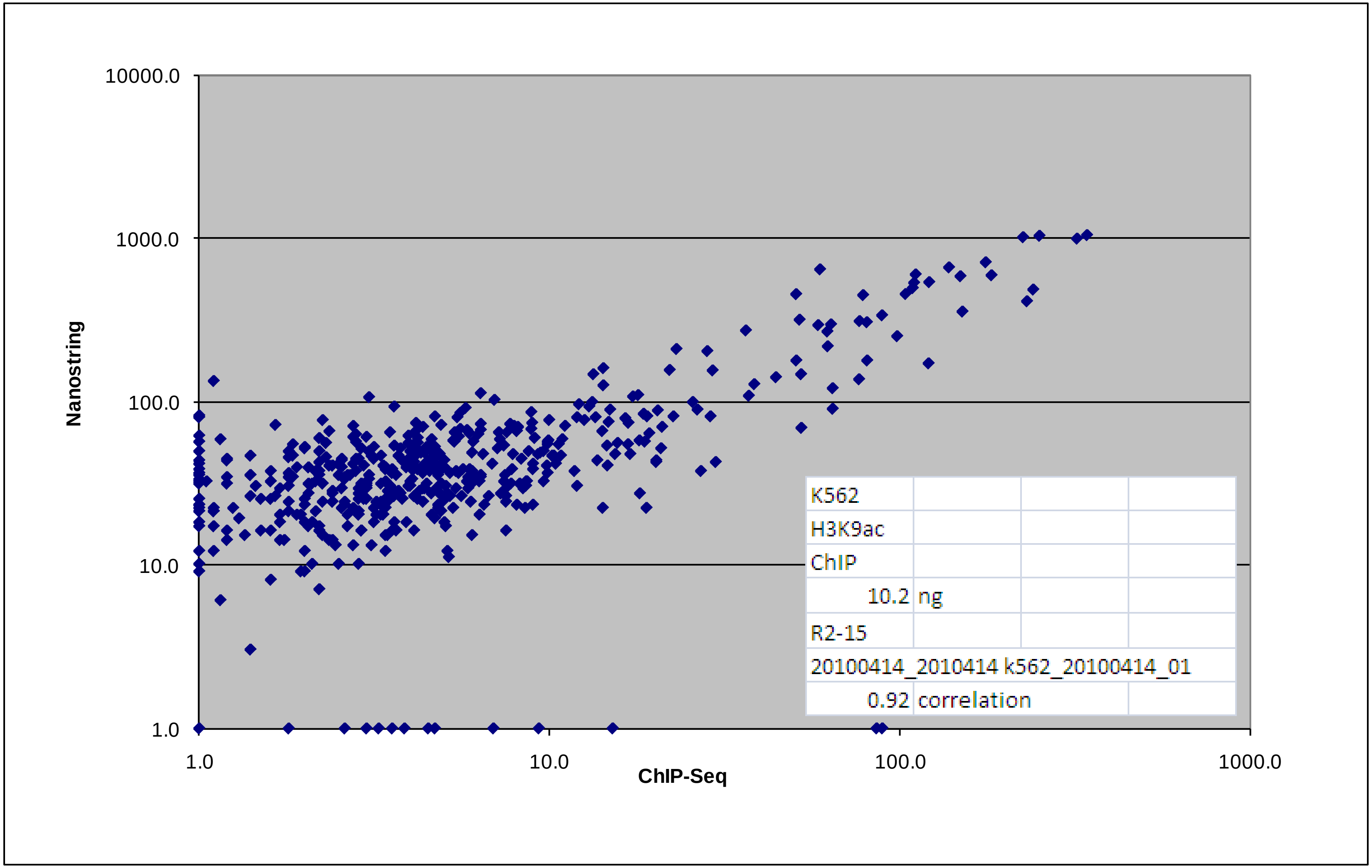
printed 1/3/2012

# K562 H3K9ac ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

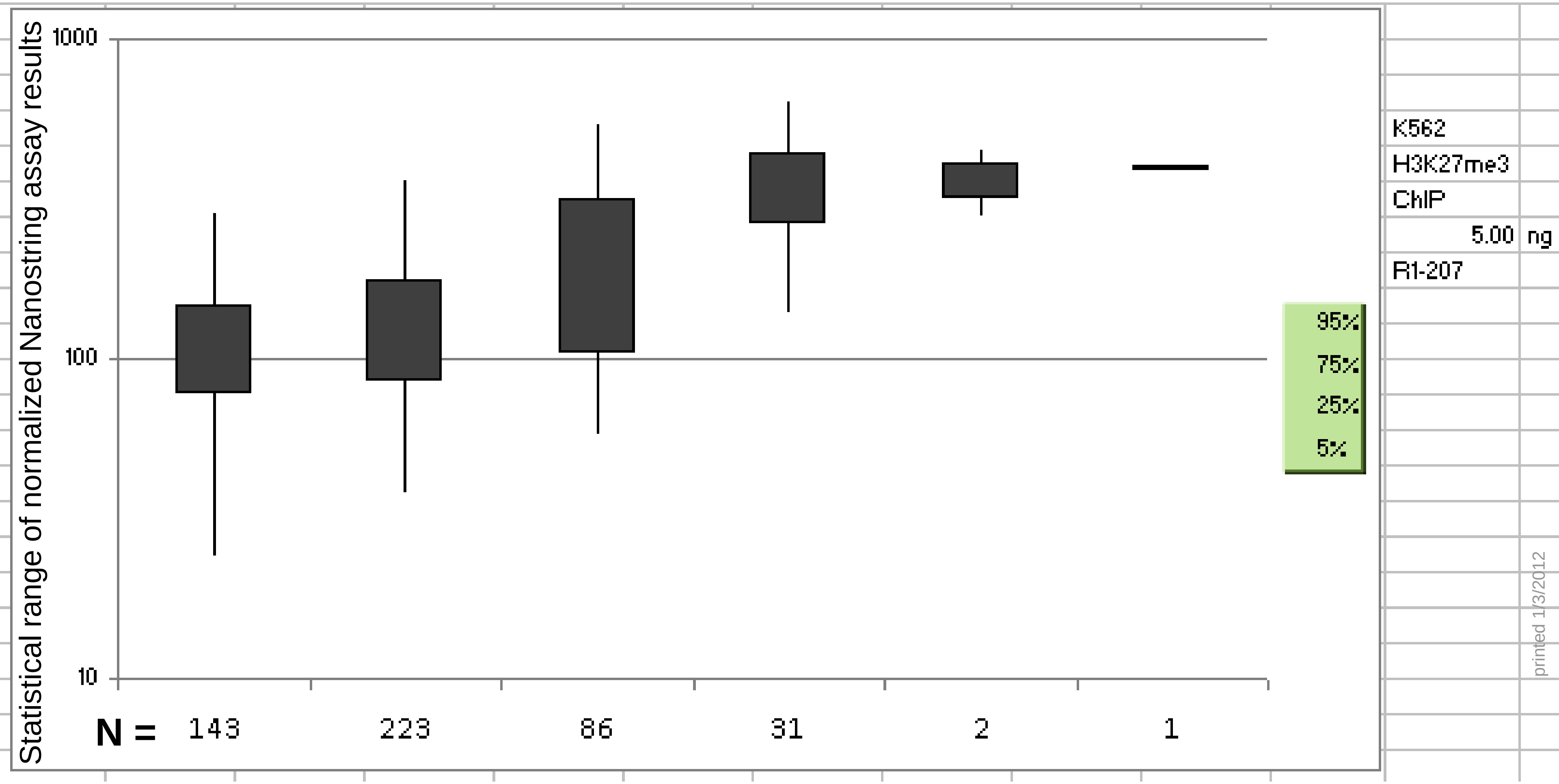
# K562 H3K9ac ChIP



| | | | |
|---|---|---|---|
| K562 | | | |
| H3K9ac | | | |
| ChIP | | | |
| 10.2 | ng | | |
| R2-15 | | | |
| 20100414_2010414 k562_20100414_01 | | | |
| 0.92 | correlation | | |

printed 1/3/2012

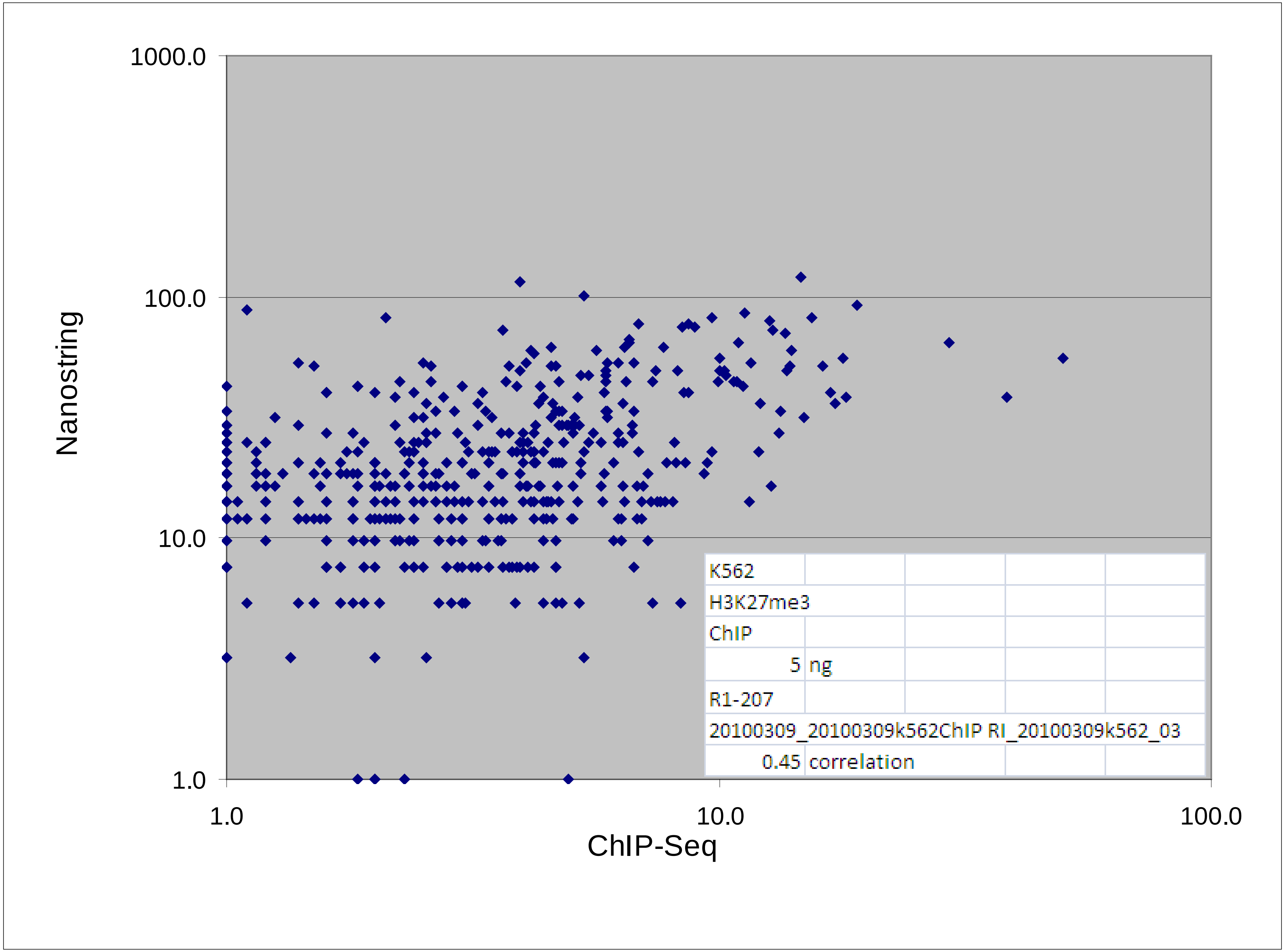
# K562 H3K27me3 ChIP



printed 1/3/2012

**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

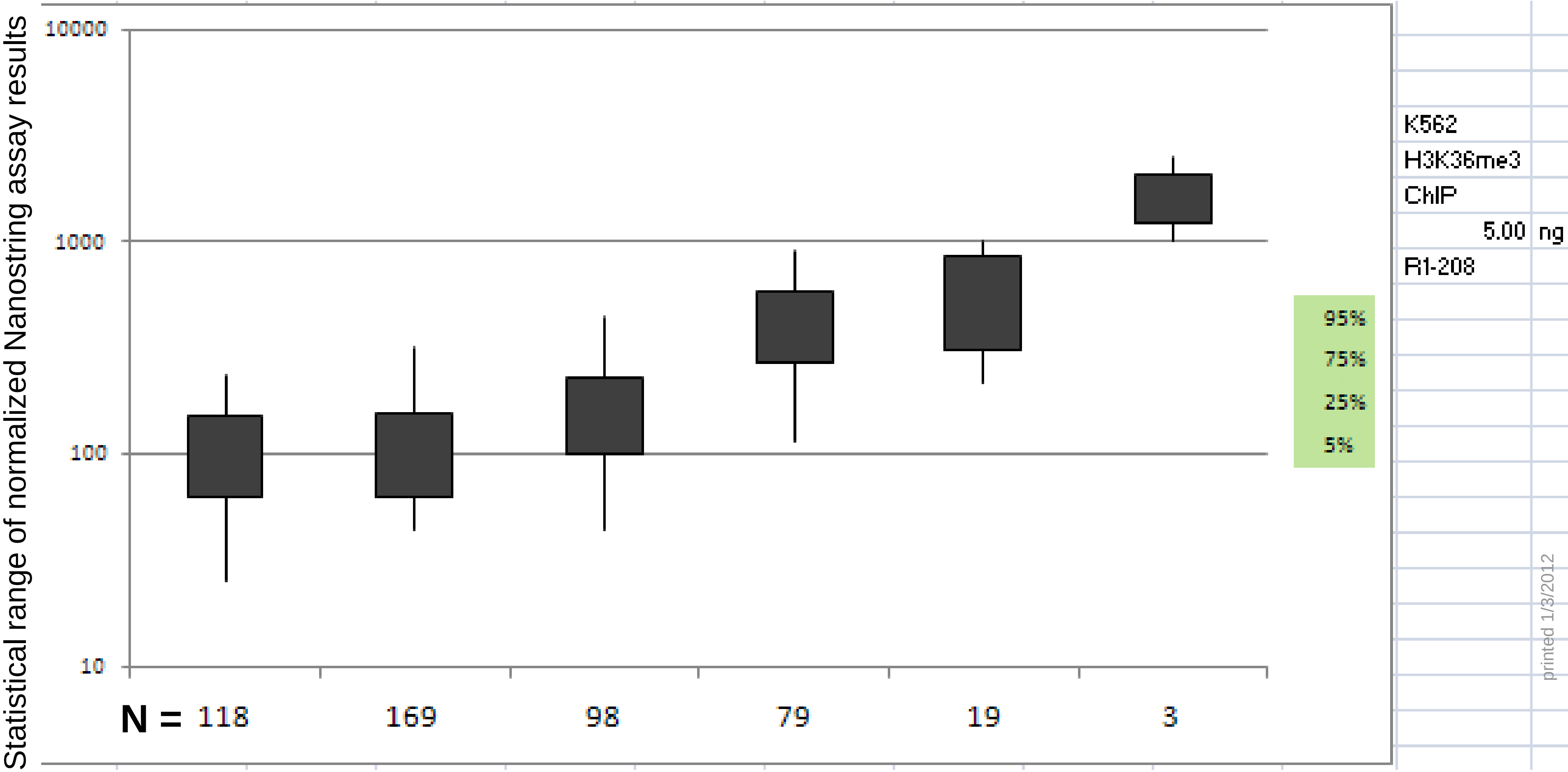
# K562 H3K27me3 ChIP



| | | | | |
|---|---|---|---|---|
| K562 | | | | |
| H3K27me3 | | | | |
| ChIP | | | | |
| 5 | ng | | | |
| R1-207 | | | | |
| 20100309_20100309k562ChIP RI_20100309k562_03 | | | | |
| 0.45 | correlation | | | |

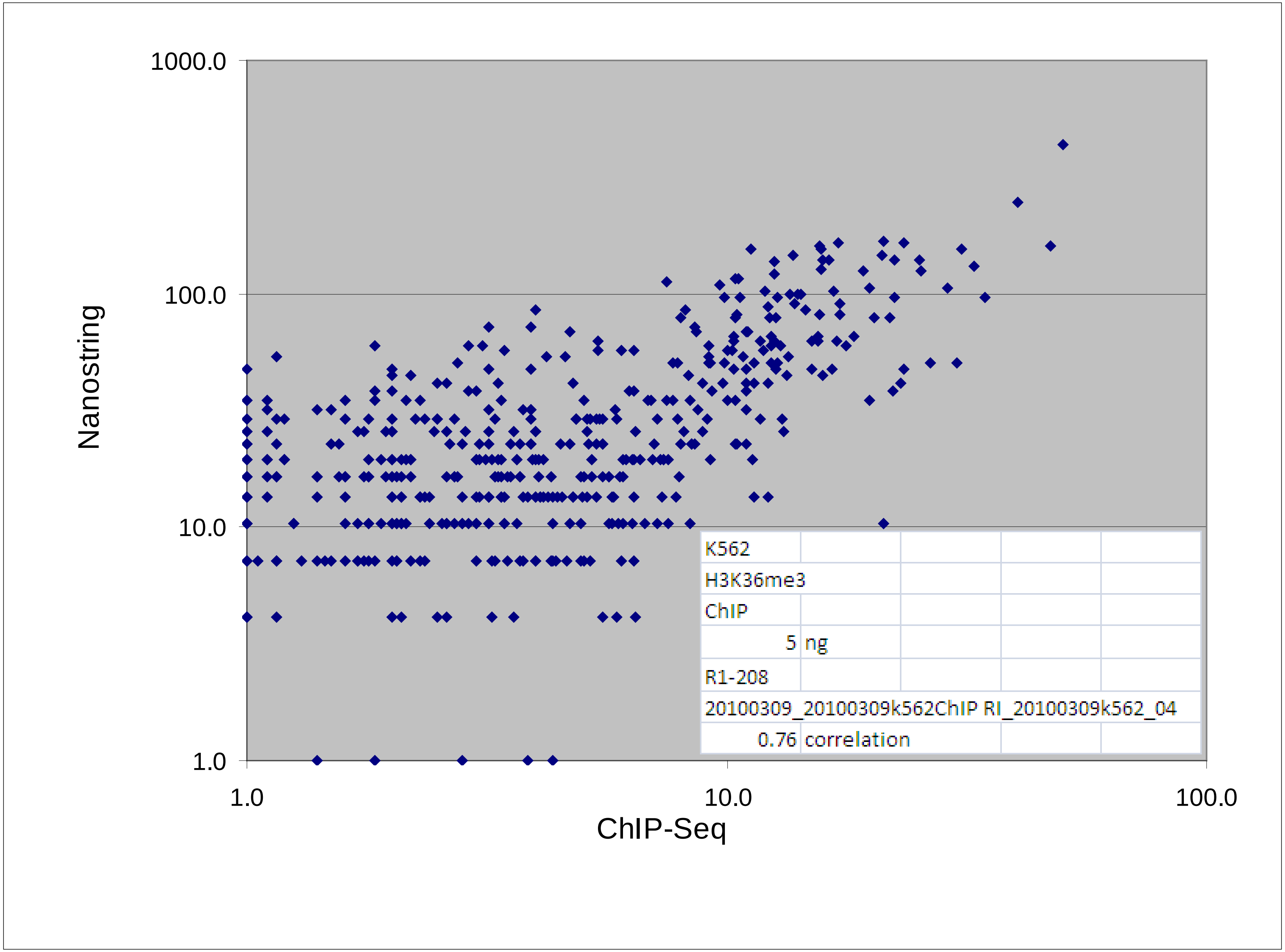
printed 1/3/2012

# K562 H3K36me3 ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2  2 – 5  5 – 10  10 – 20  20 – 40  > 40**

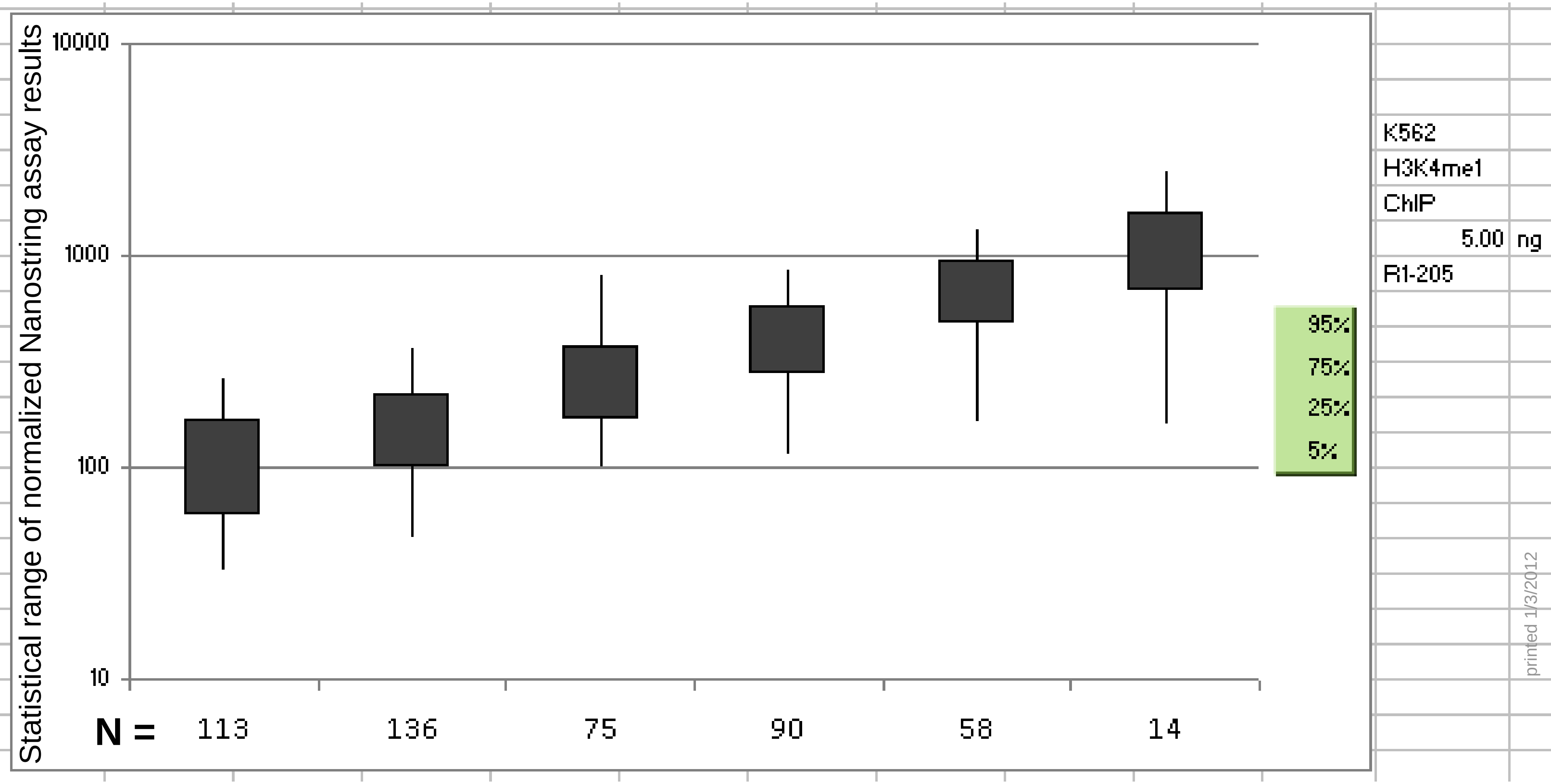
# K562 H3K36me3 ChIP



| | |
|---|---|
| K562 | |
| H3K36me3 | |
| ChIP | |
| 5 | ng |
| R1-208 | |
| 20100309_20100309k562ChIP RI_20100309k562_04 | |
| 0.76 | correlation |

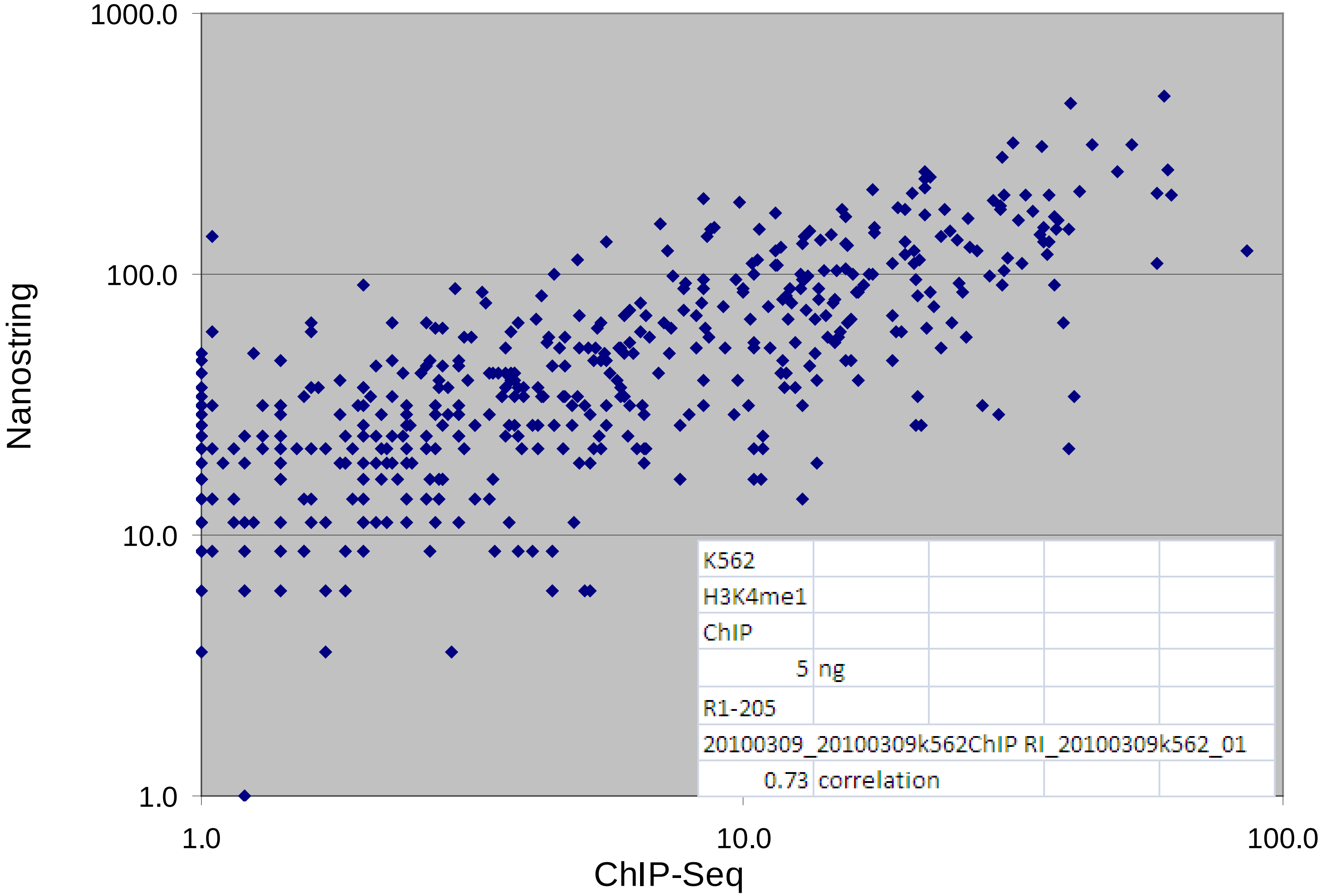
printed 1/3/2012

# K562 H3K4me1 ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

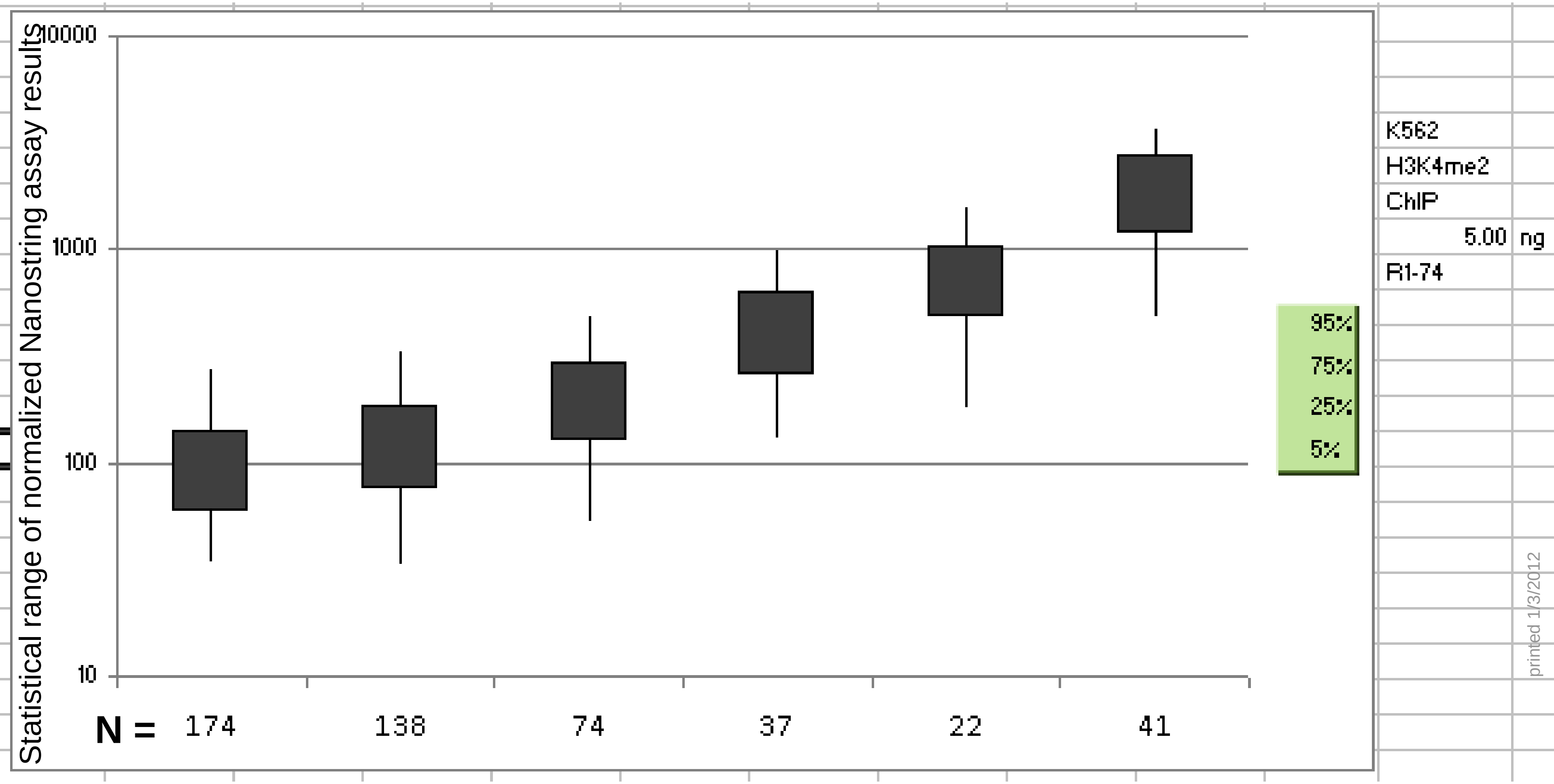
# K562 H3K4me1 ChIP



| | | | | |
|---|---|---|---|---|
| K562 | | | | |
| H3K4me1 | | | | |
| ChIP | | | | |
| 5 | ng | | | |
| R1-205 | | | | |
| 20100309_20100309k562ChIP RI_20100309k562_01 | | | | |
| 0.73 | correlation | | | |

printed 1/3/2012

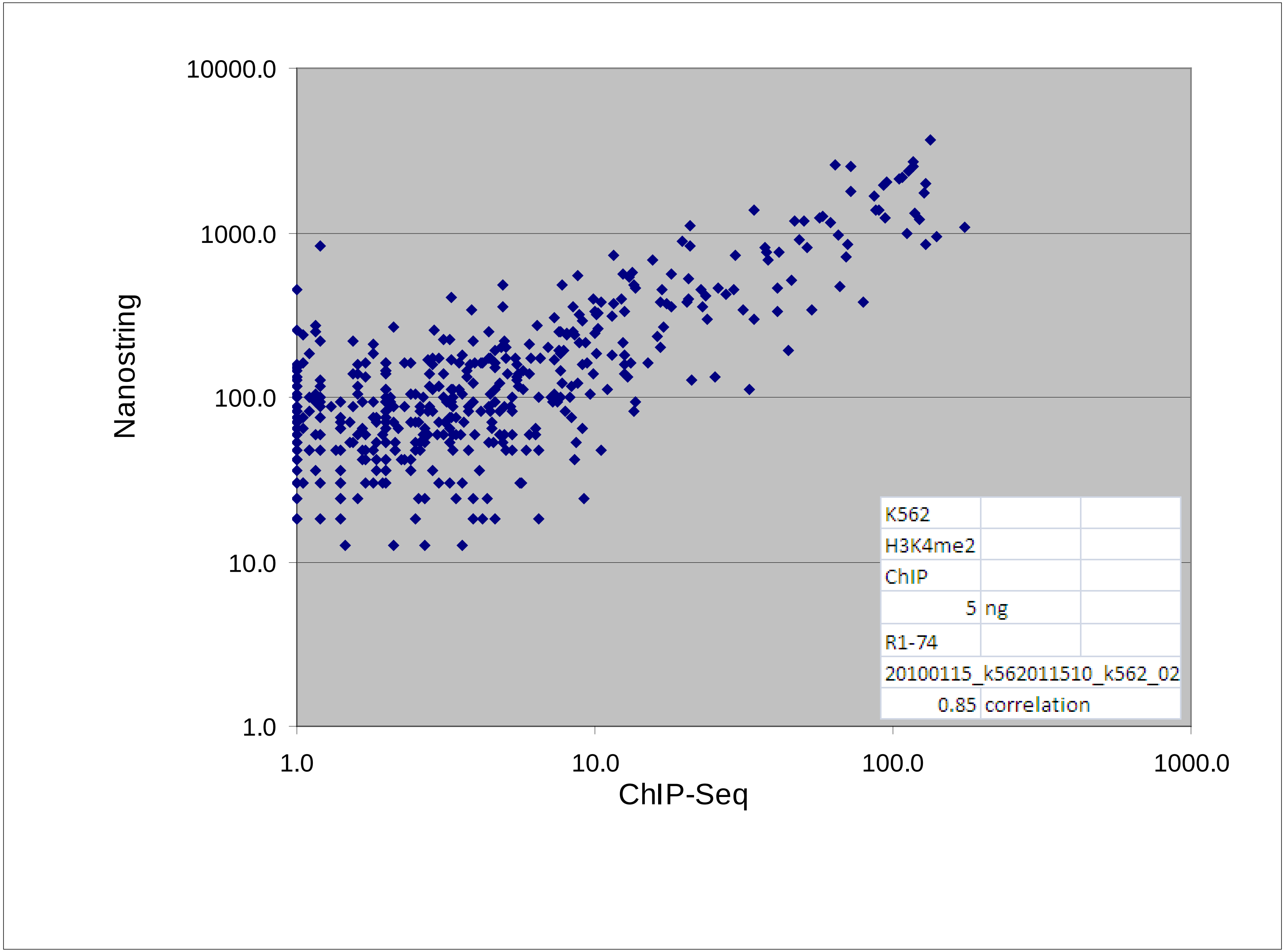
# K562 H3K4me2 ChIP



Statistical range of normalized Nanostring assay results

**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

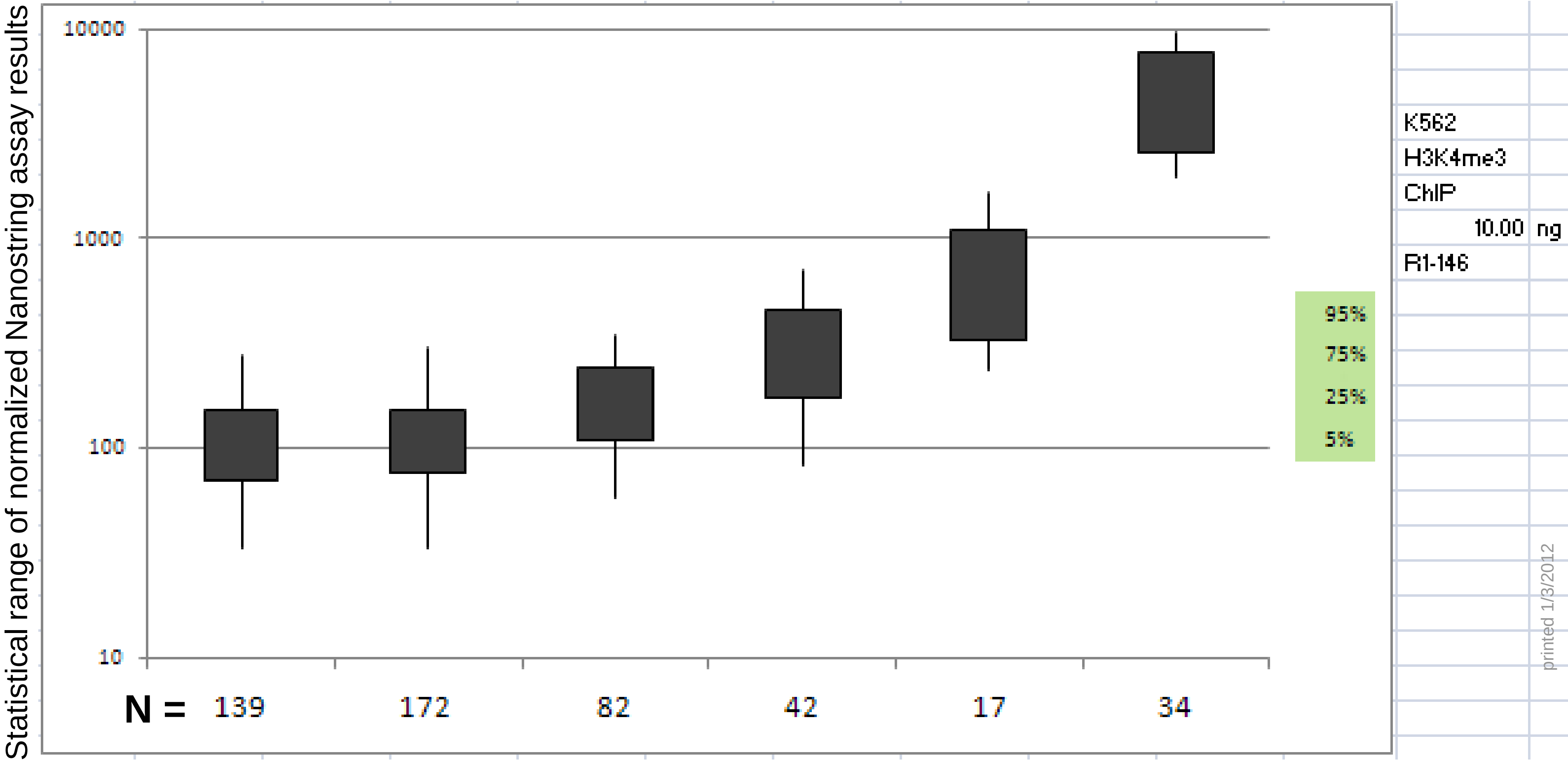
# K562 H3K4me2 ChIP



| | | |
|---|---|---|
| K562 | | |
| H3K4me2 | | |
| ChIP | | |
| 5 | ng | |
| R1-74 | | |
| 20100115_k562011510_k562_02 | | |
| 0.85 | correlation | |

printed 1/3/2012

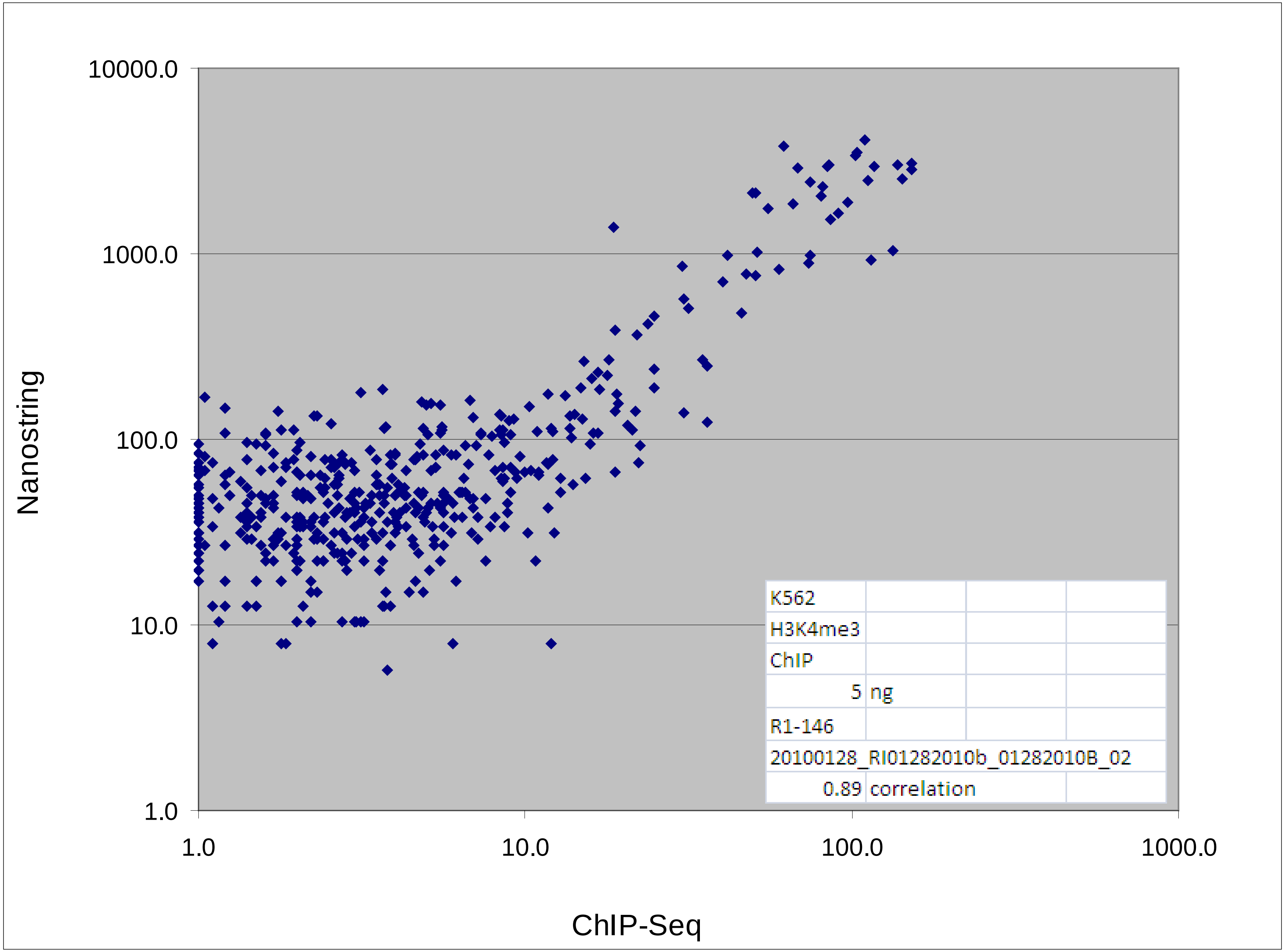
# K562 H3K4me3 ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**

**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

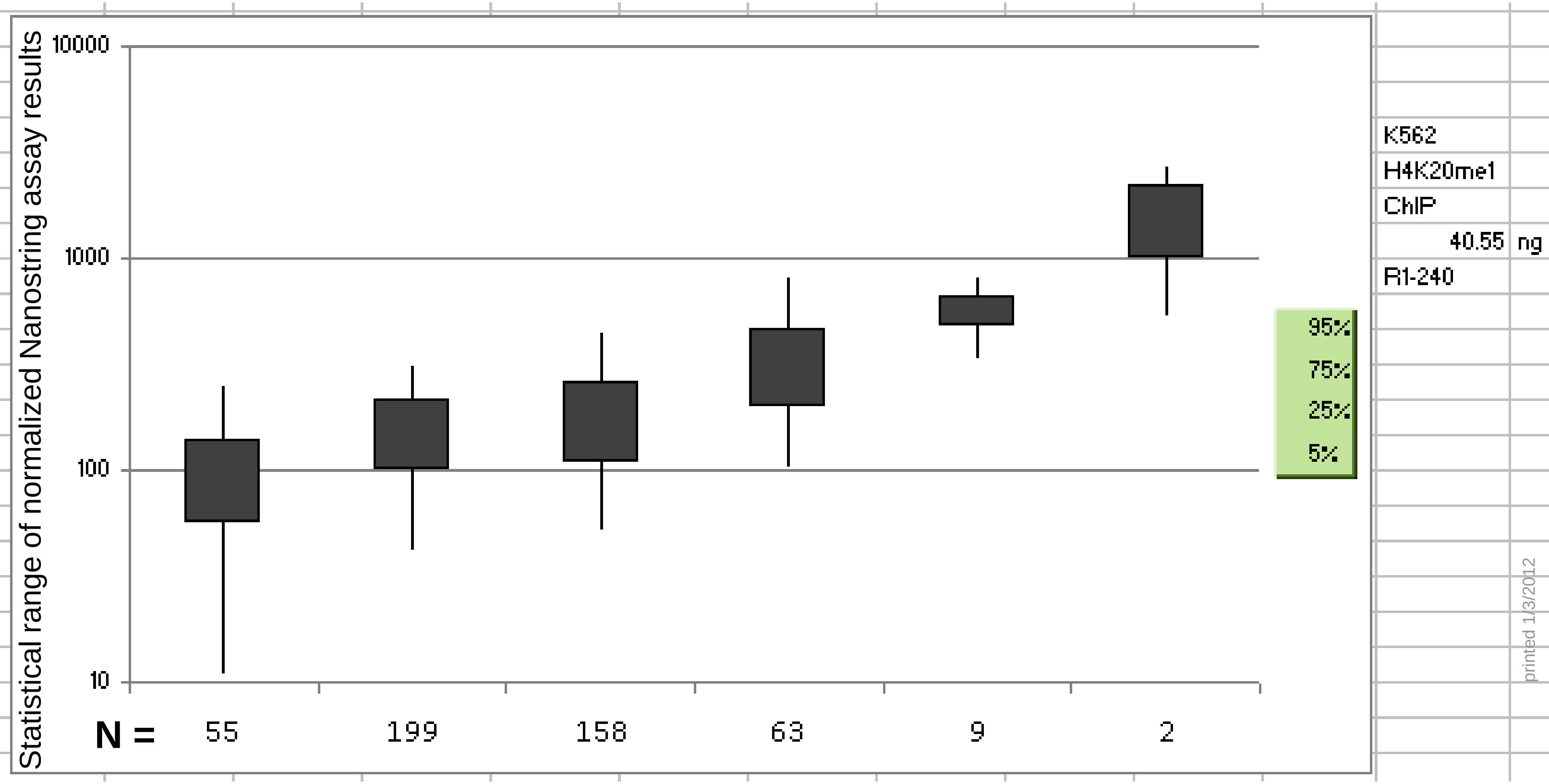
# K562 H3K4me3 ChIP



| | | | |
|---|---|---|---|
| K562 | | | |
| H3K4me3 | | | |
| ChIP | | | |
| 5 | ng | | |
| R1-146 | | | |
| 20100128_RI01282010b_01282010B_02 | | | |
| 0.89 | correlation | | |

printed 1/3/2012

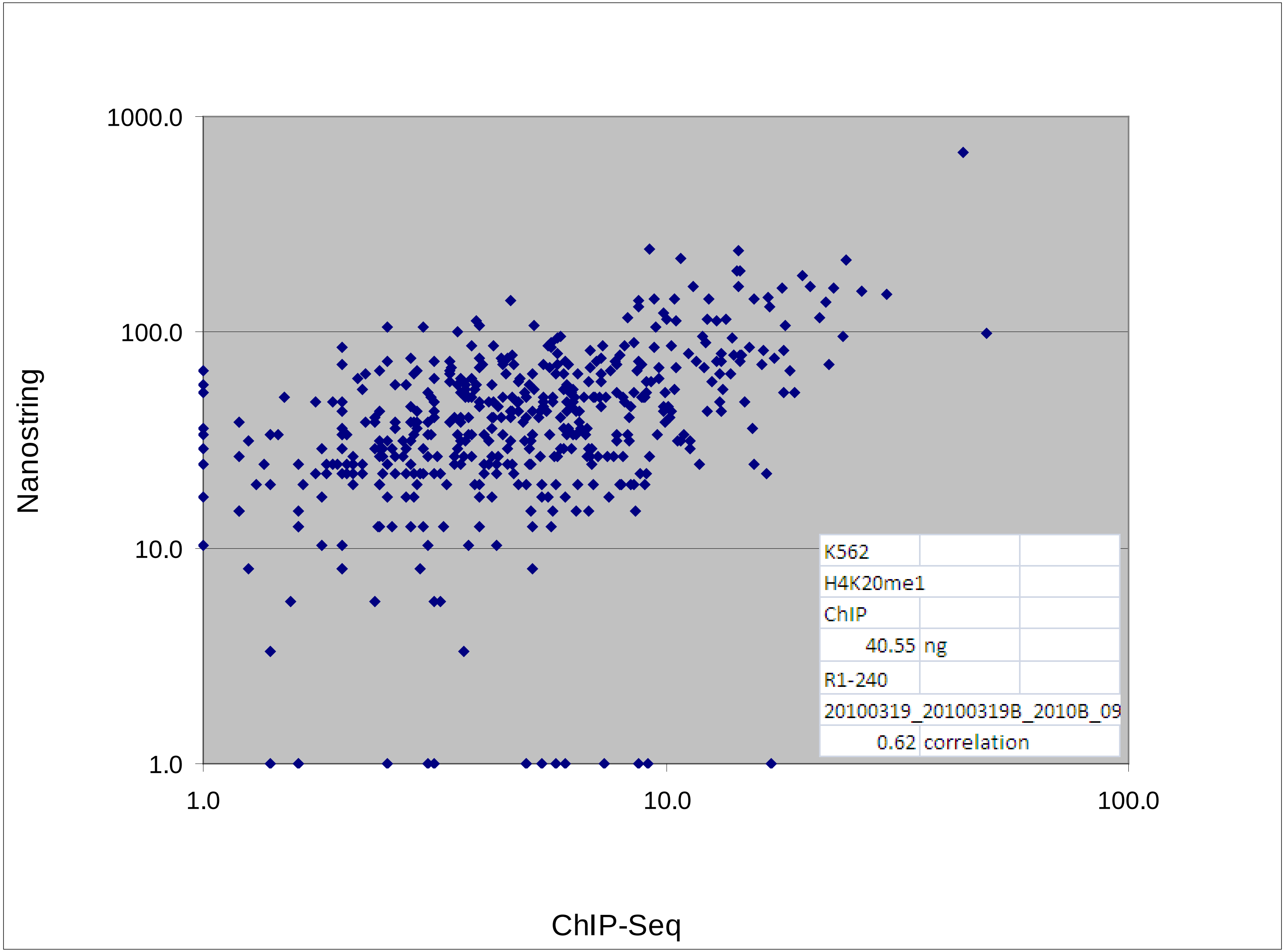
# K562 H4K20me1 ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**

**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

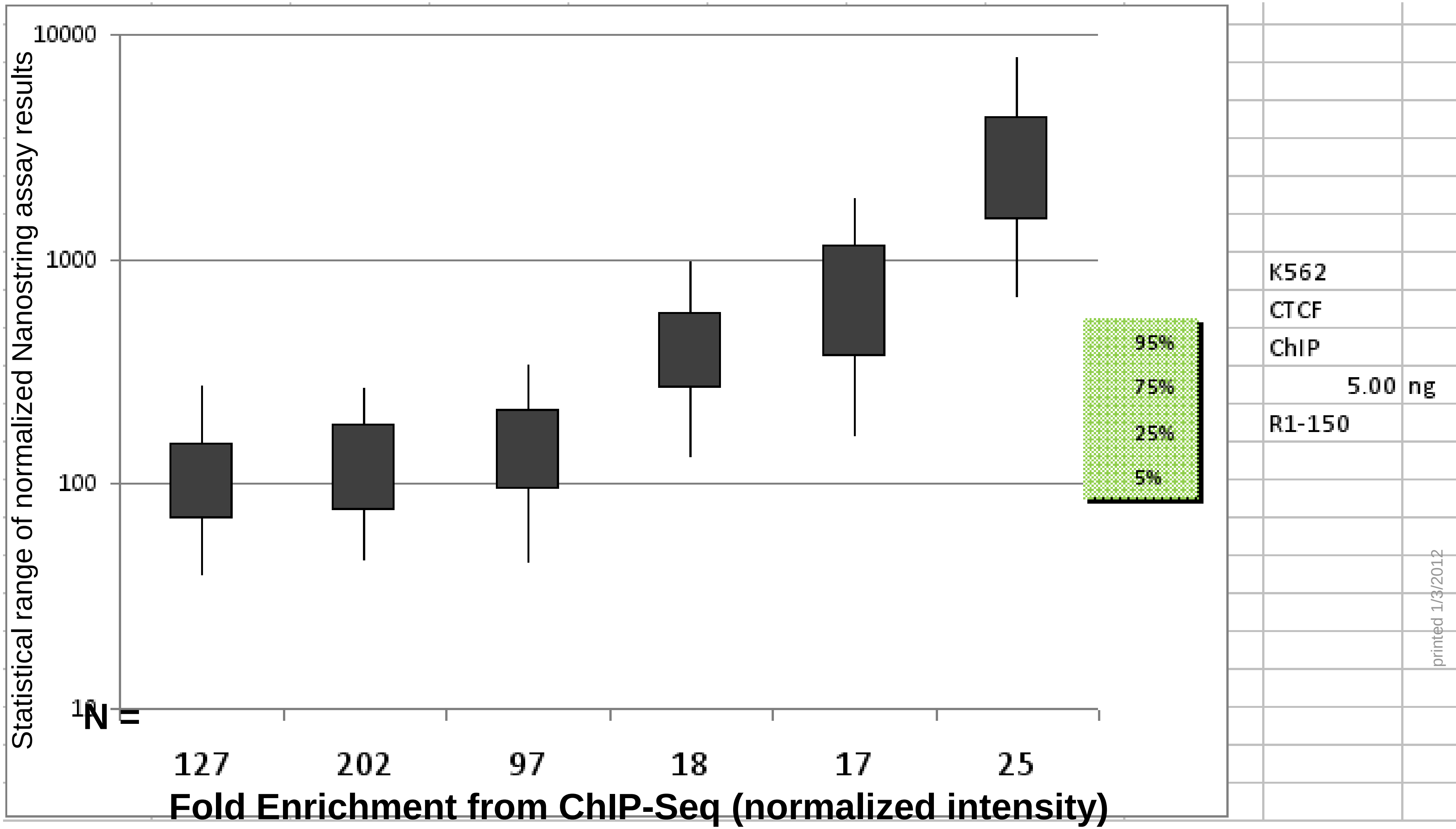
# K562 H4K20me1 ChIP



| | | |
|---|---|---|
| K562 | | |
| H4K20me1 | | |
| ChIP | | |
| 40.55 | ng | |
| R1-240 | | |
| 20100319_20100319B_2010B_09 | | |
| 0.62 | correlation | |

printed 1/3/2012

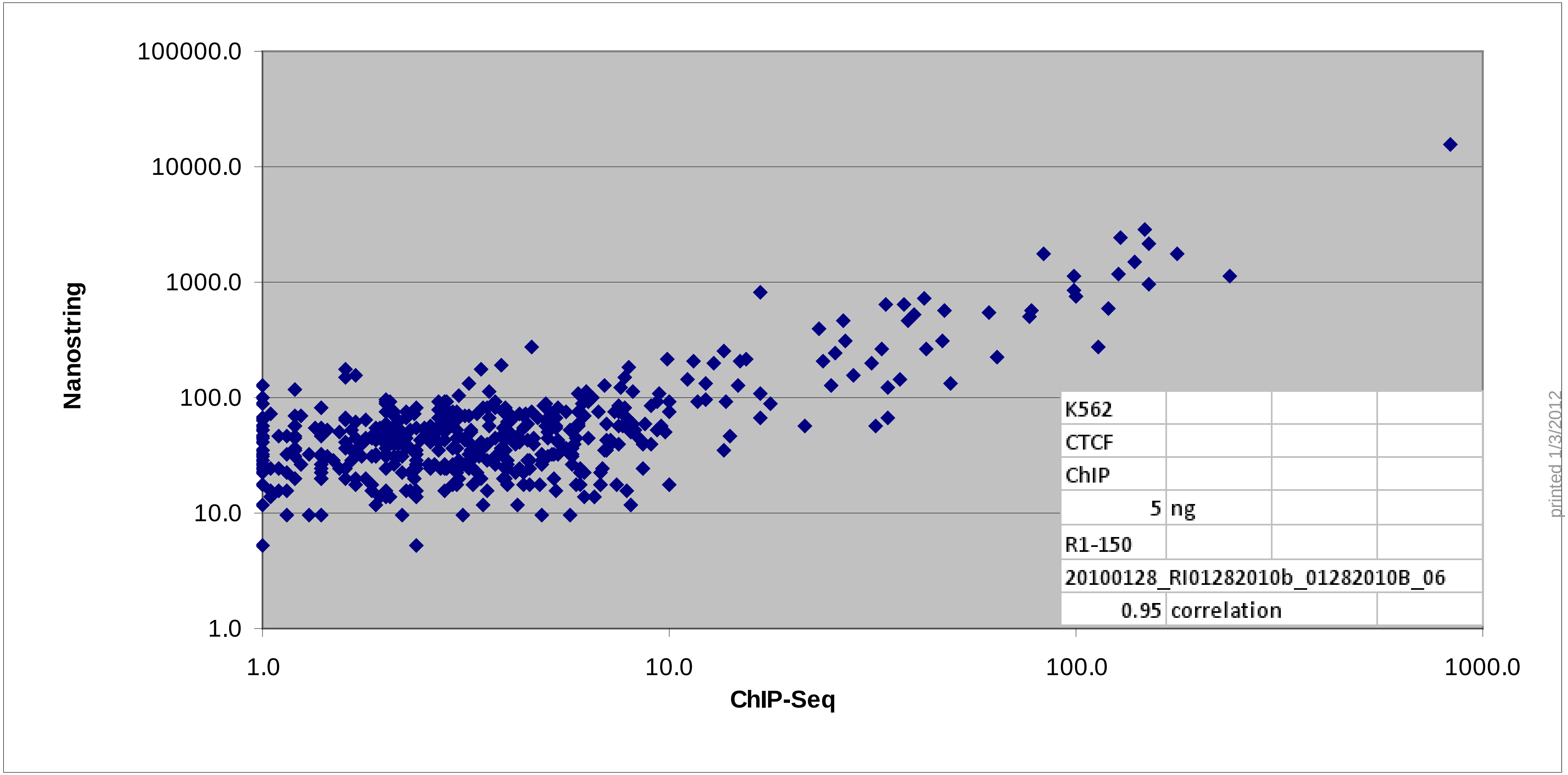
# K562 CTCF Library



**Fold Enrichment from ChIP-Seq (normalized intensity)**

**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

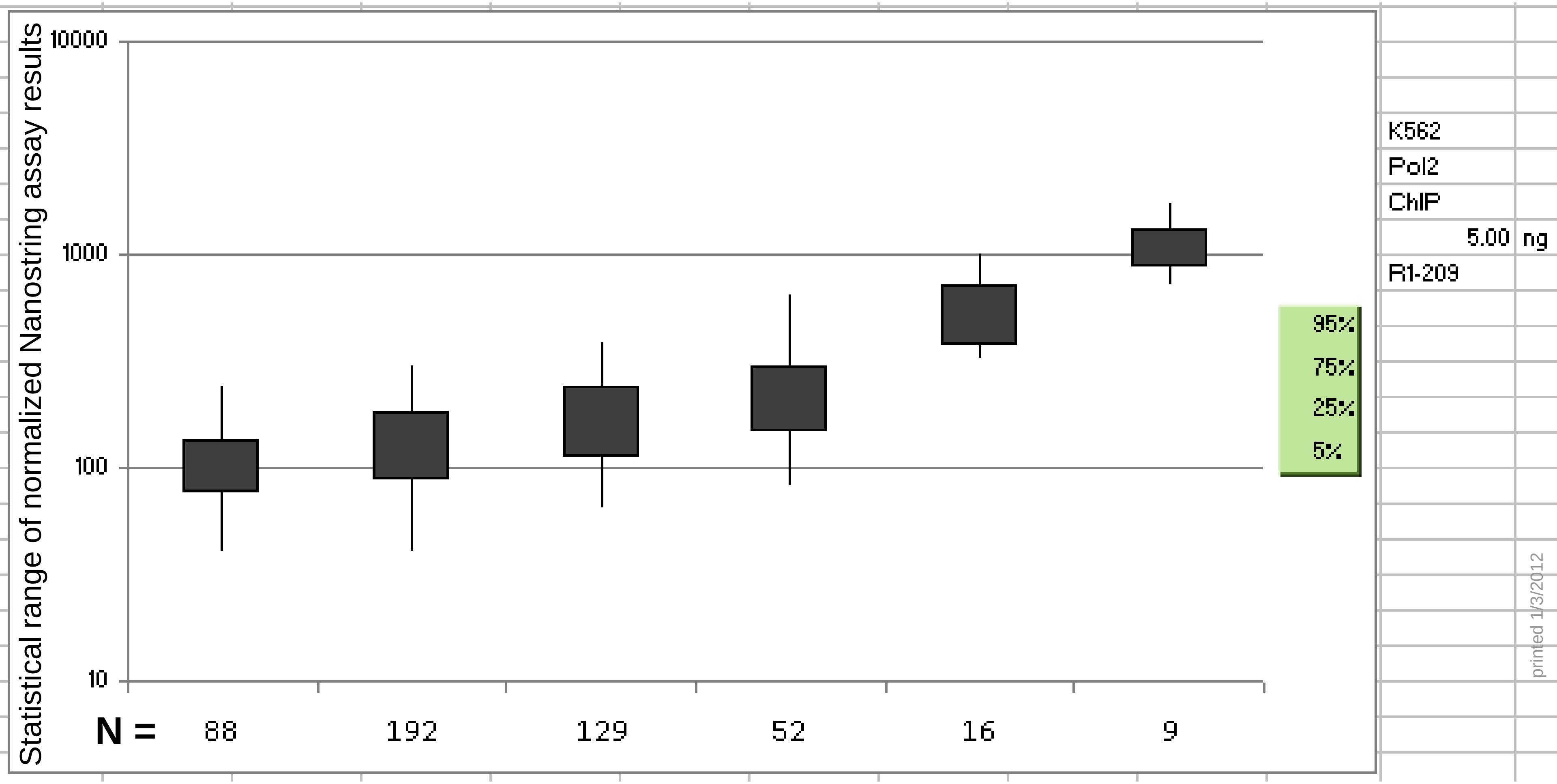
# K562 CTCF Library



| | | | |
|---|---|---|---|
| K562 | | | |
| CTCF | | | |
| ChIP | | | |
| 5 | ng | | |
| R1-150 | | | |
| 20100128_RI01282010b_01282010B_06 | | | |
| 0.95 | correlation | | |

printed 1/3/2012

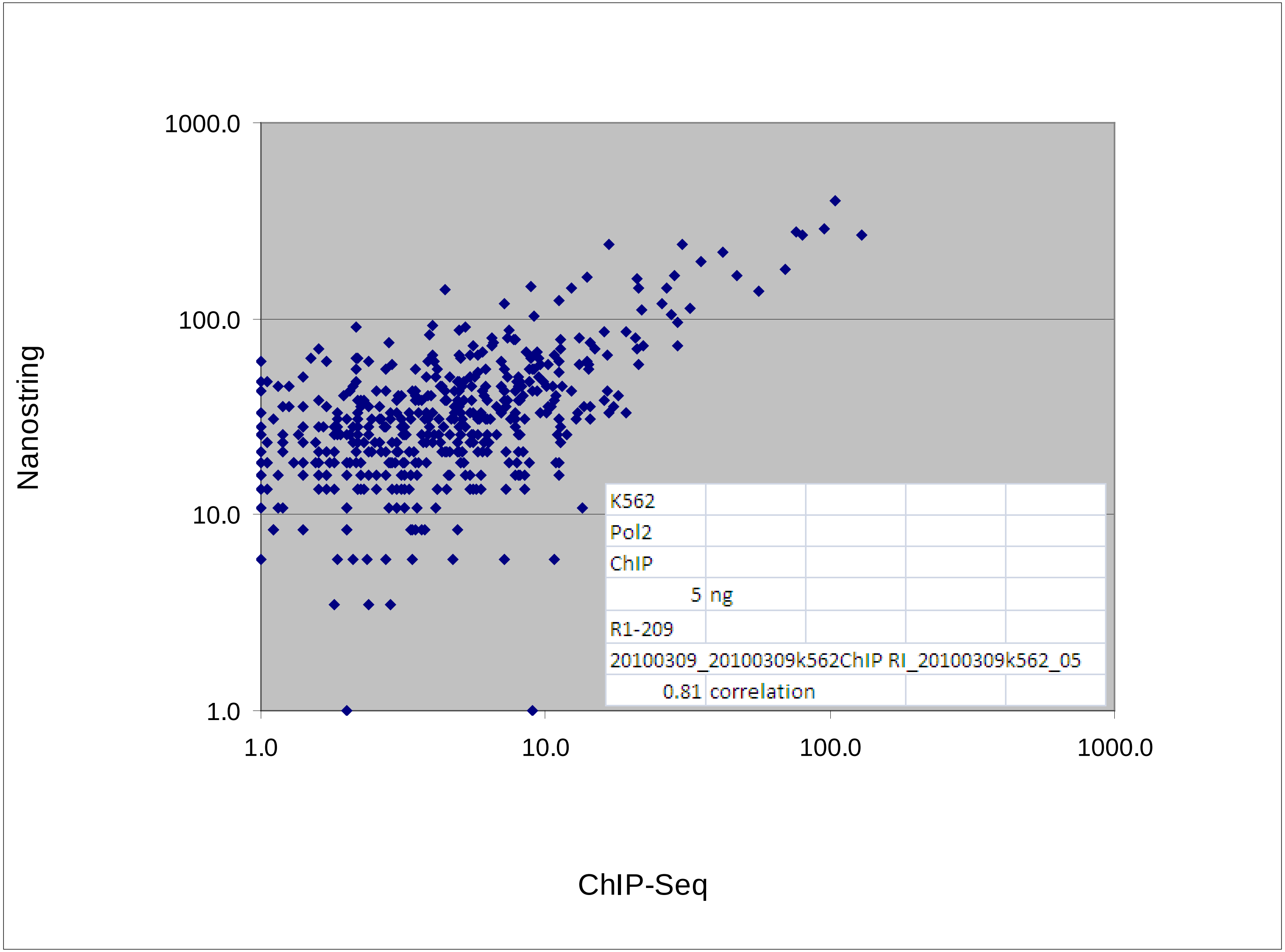
# K562 RNA Pol II ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**

**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

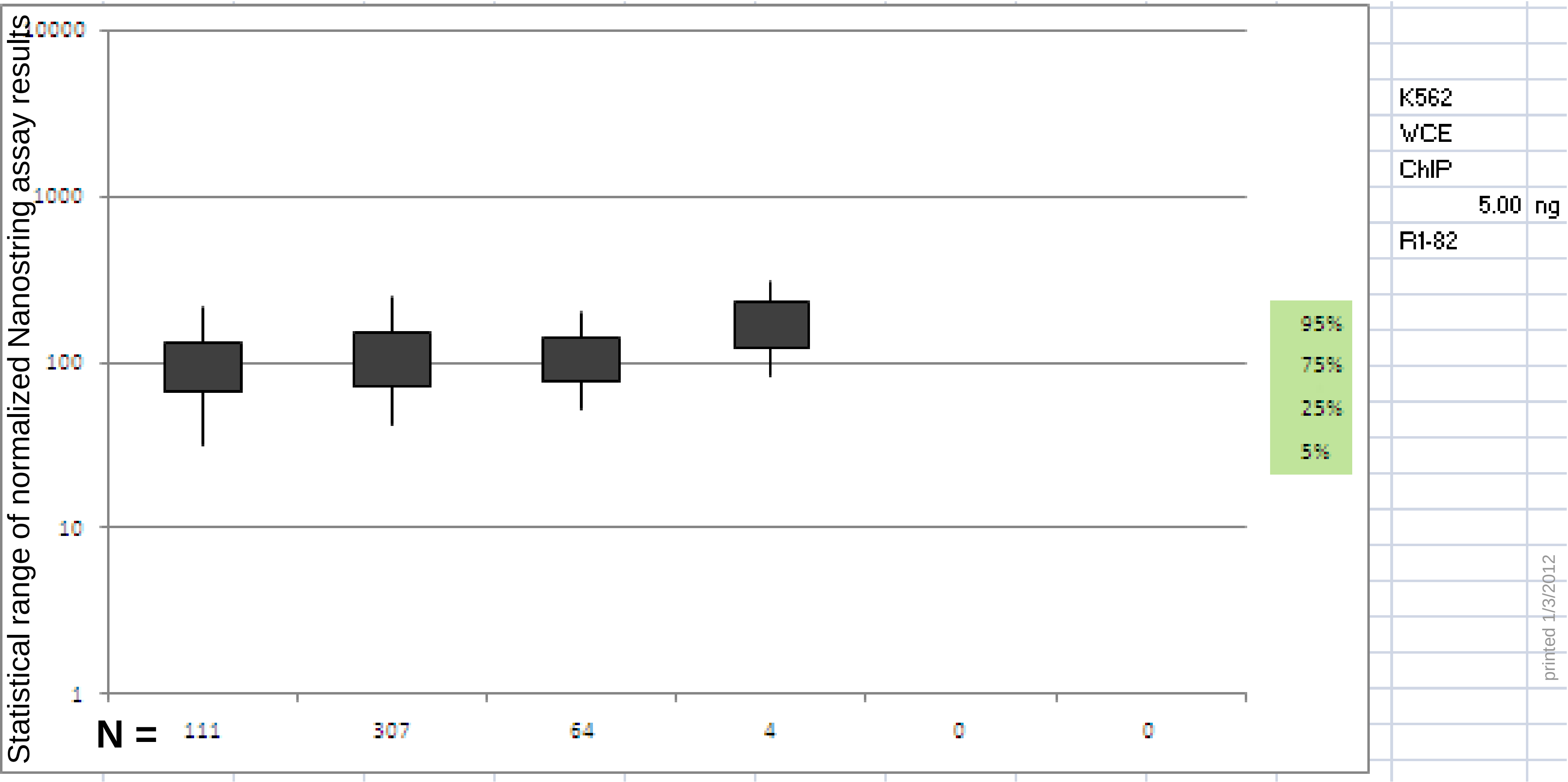
# K562 RNA Pol II ChIP



| | | | | |
|---|---|---|---|---|
| K562 | | | | |
| Pol2 | | | | |
| ChIP | | | | |
| 5 | ng | | | |
| R1-209 | | | | |
| 20100309_20100309k562ChIP RI_20100309k562_05 | | | | |
| 0.81 | correlation | | | |

printed 1/3/2012

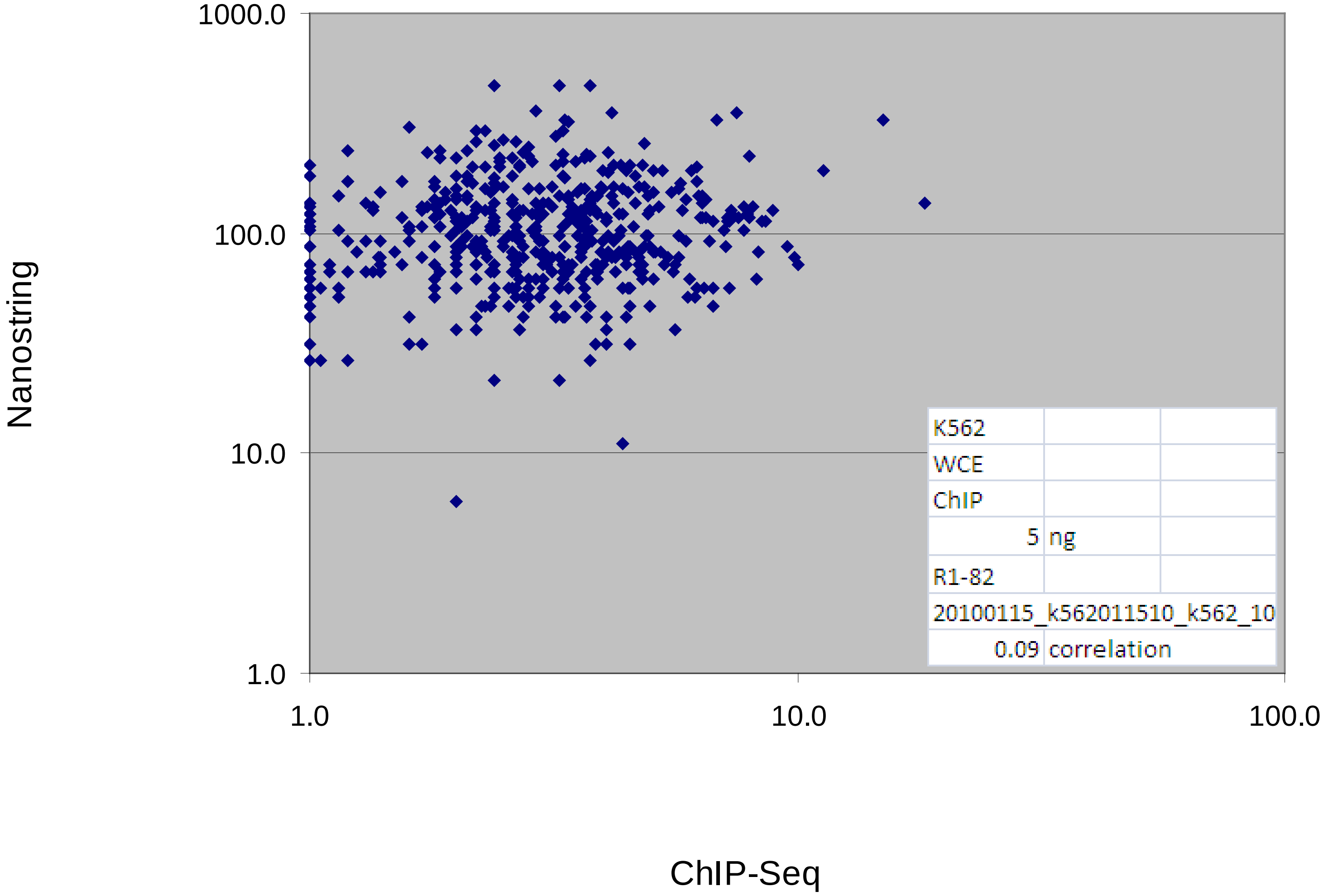
# K562 whole cell extract – chromatin prep



**Fold Enrichment from ChIP-Seq (normalized intensity)**

**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

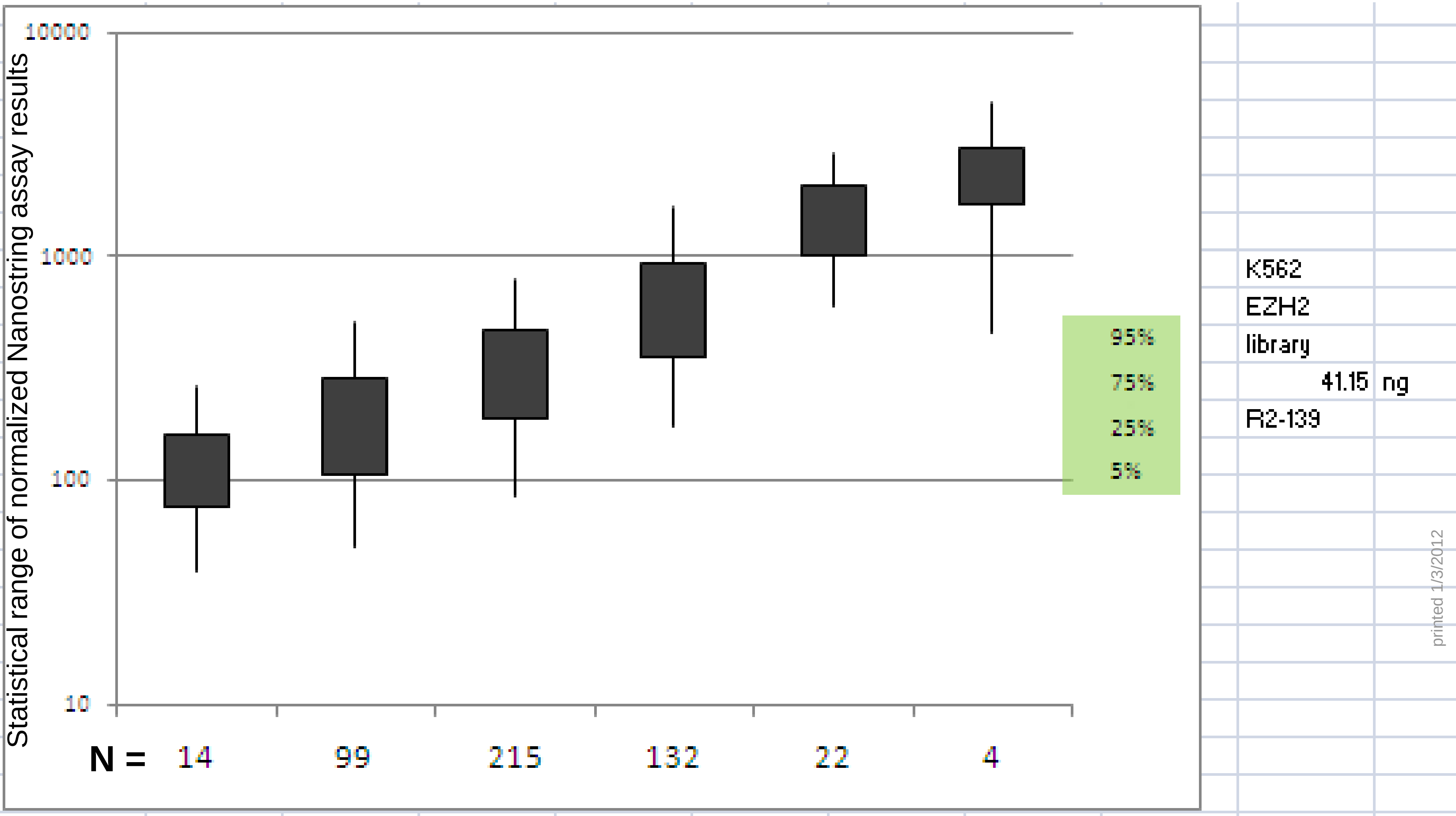
# K562 whole cell extract – chromatin prep



| | | |
|---|---|---|
| K562 | | |
| WCE | | |
| ChIP | | |
| 5 | ng | |
| R1-82 | | |
| 20100115_k562011510_k562_10 | | |
| 0.09 | correlation | |

printed 1/3/2012

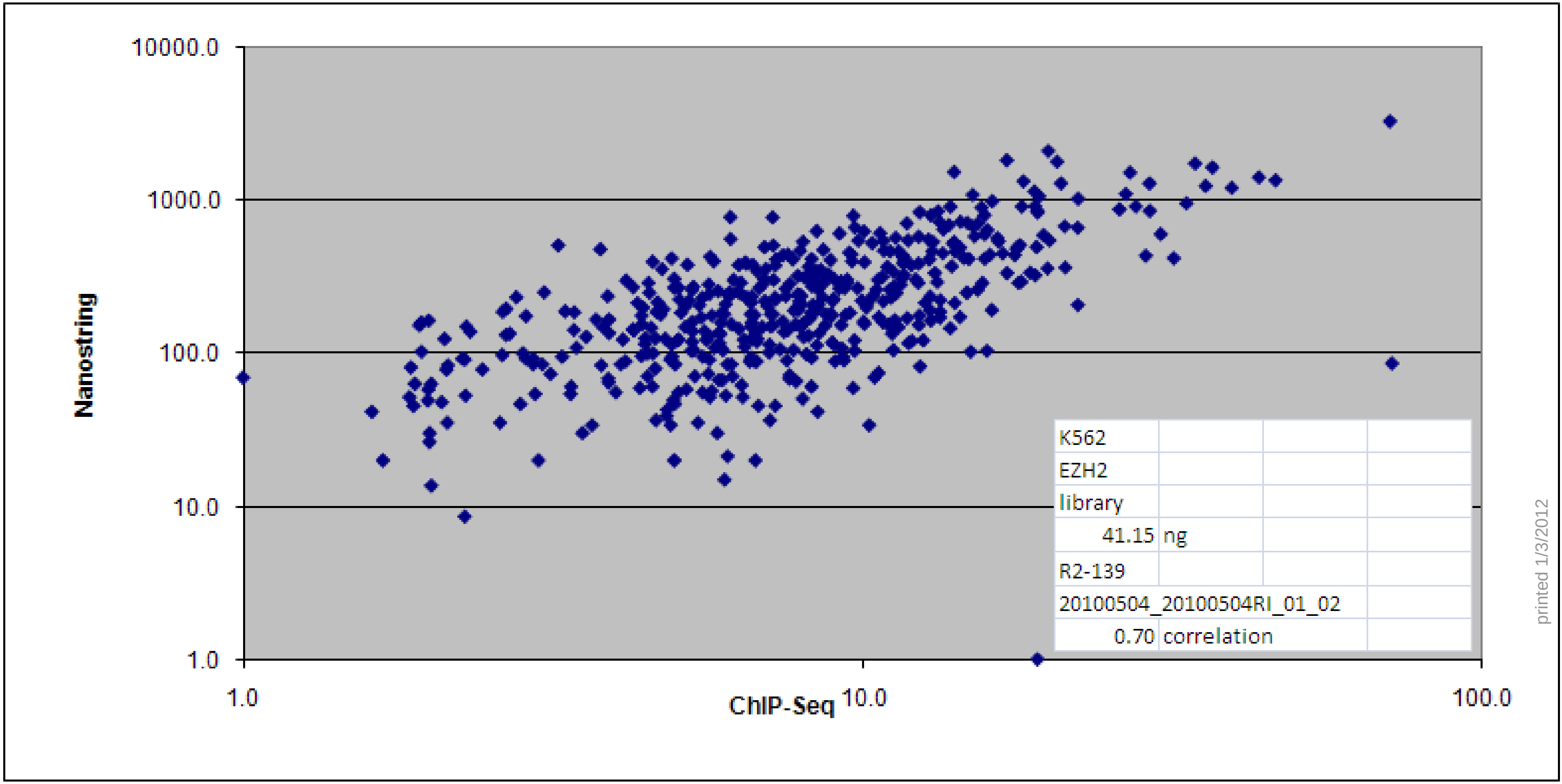
# K562 EZH2 library DNA



**Fold Enrichment from ChIP-Seq (normalized intensity)**

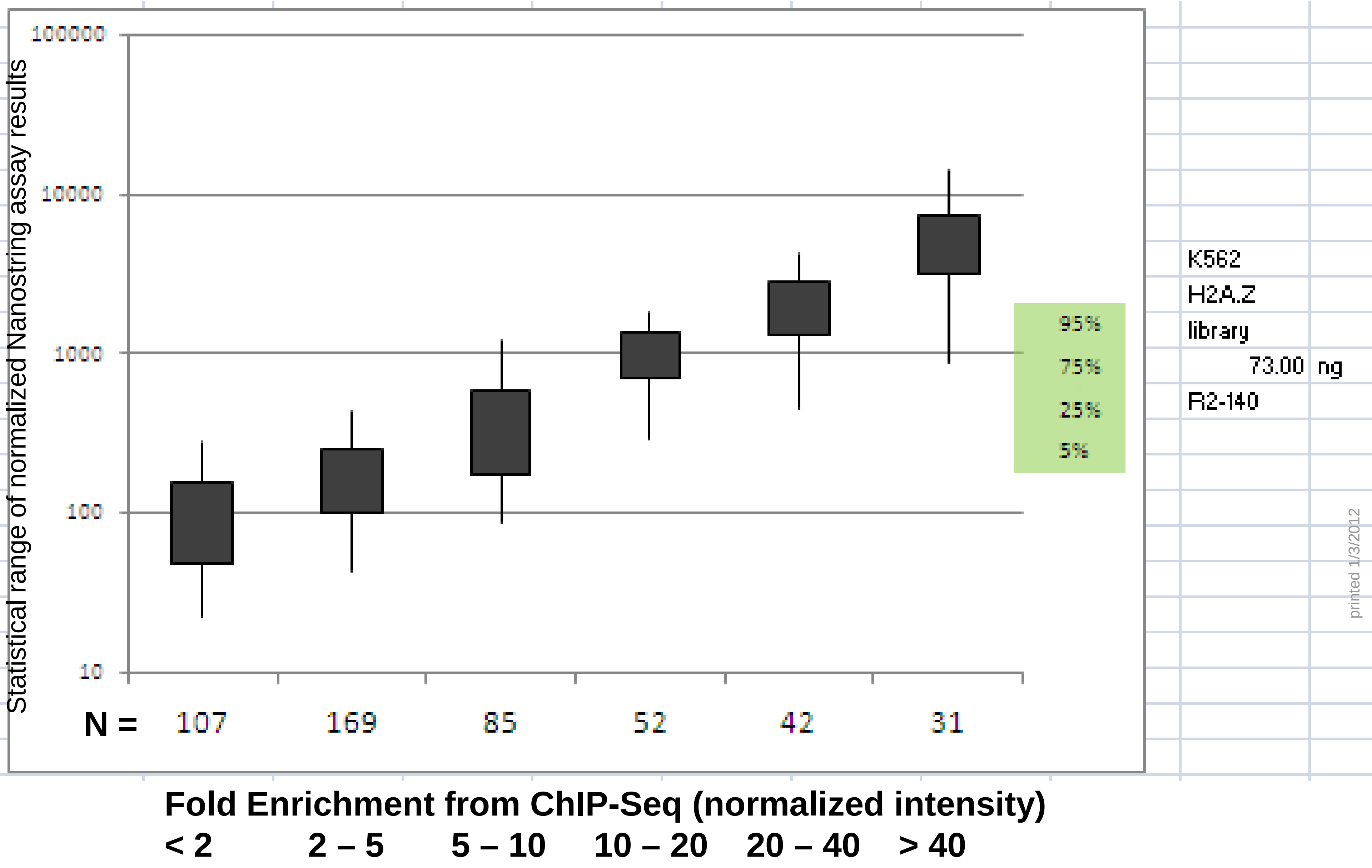
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

# K562 EZH2 library DNA

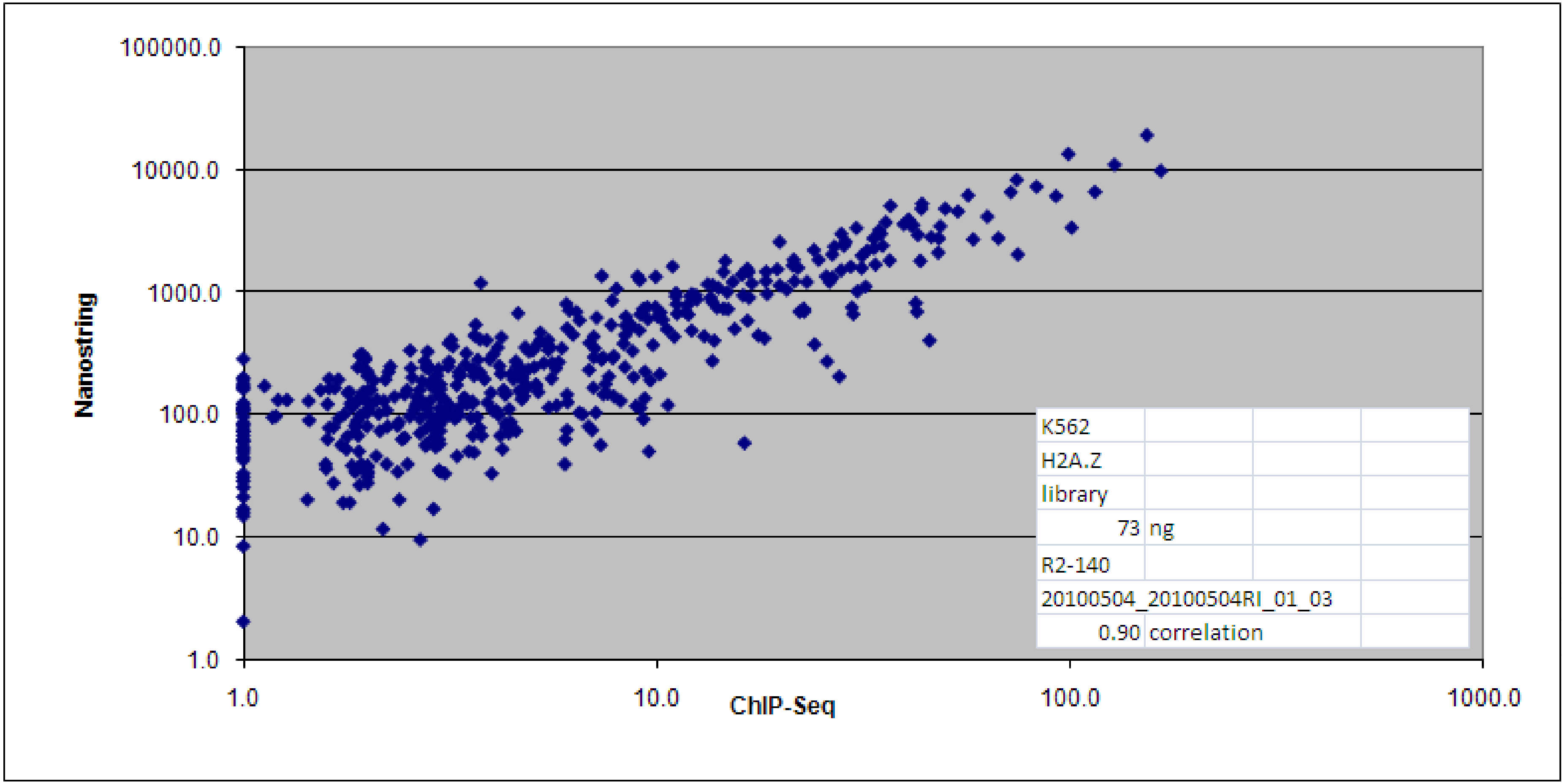


printed 1/3/2012

# K562 H2A.Z library DNA



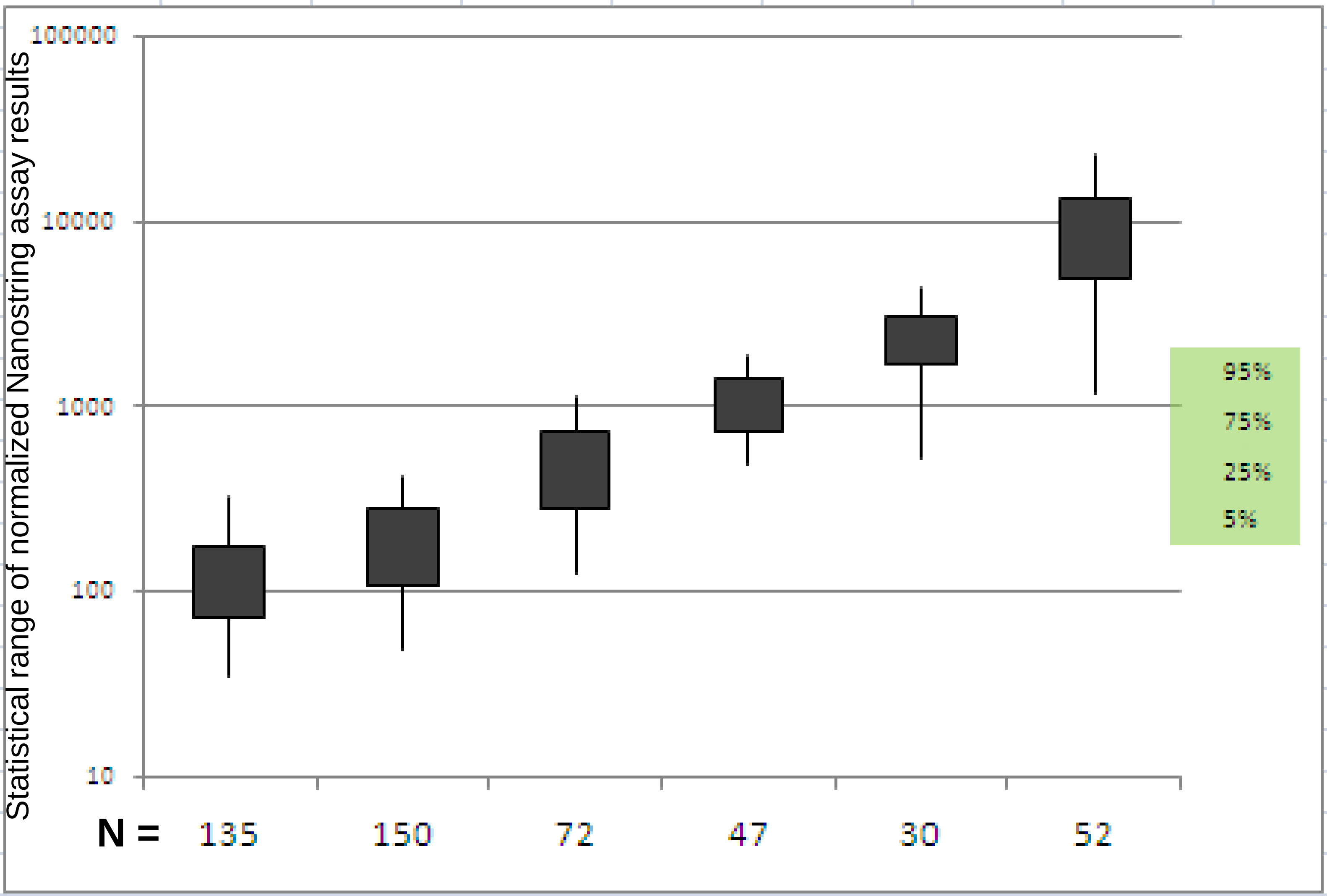
# K562 H2A.Z library DNA



| | | | |
|---|---|---|---|
| K562 | | | |
| H2A.Z | | | |
| library | | | |
| 73 | ng | | |
| R2-140 | | | |
| 20100504_20100504RI_01_03 | | | |
| 0.90 | correlation | | |

printed 1/3/2012

# K562 H3K79me2 library DNA

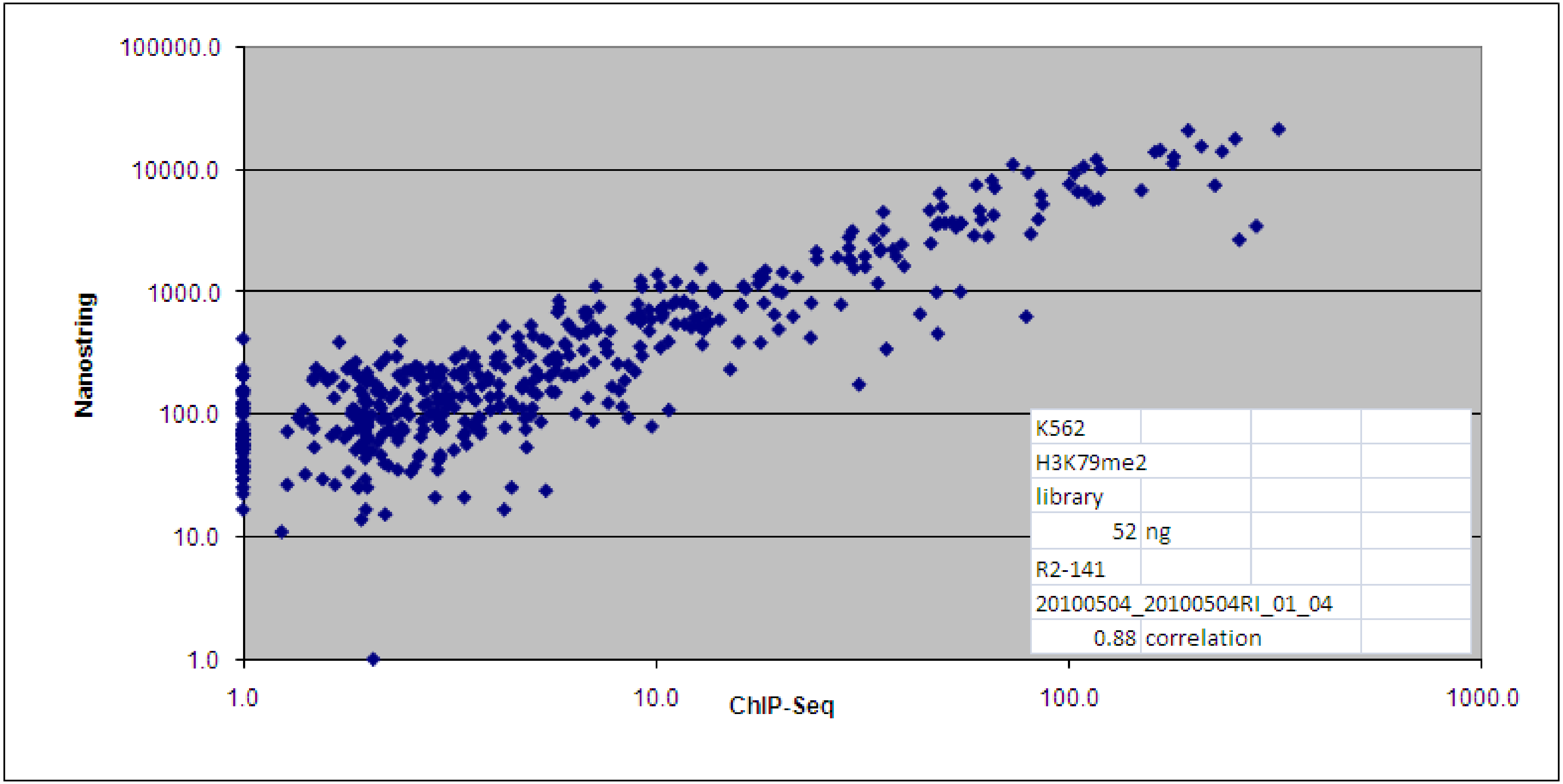


K562
H3K79me2
library
52.00 ng
R2-141

printed 1/3/2012

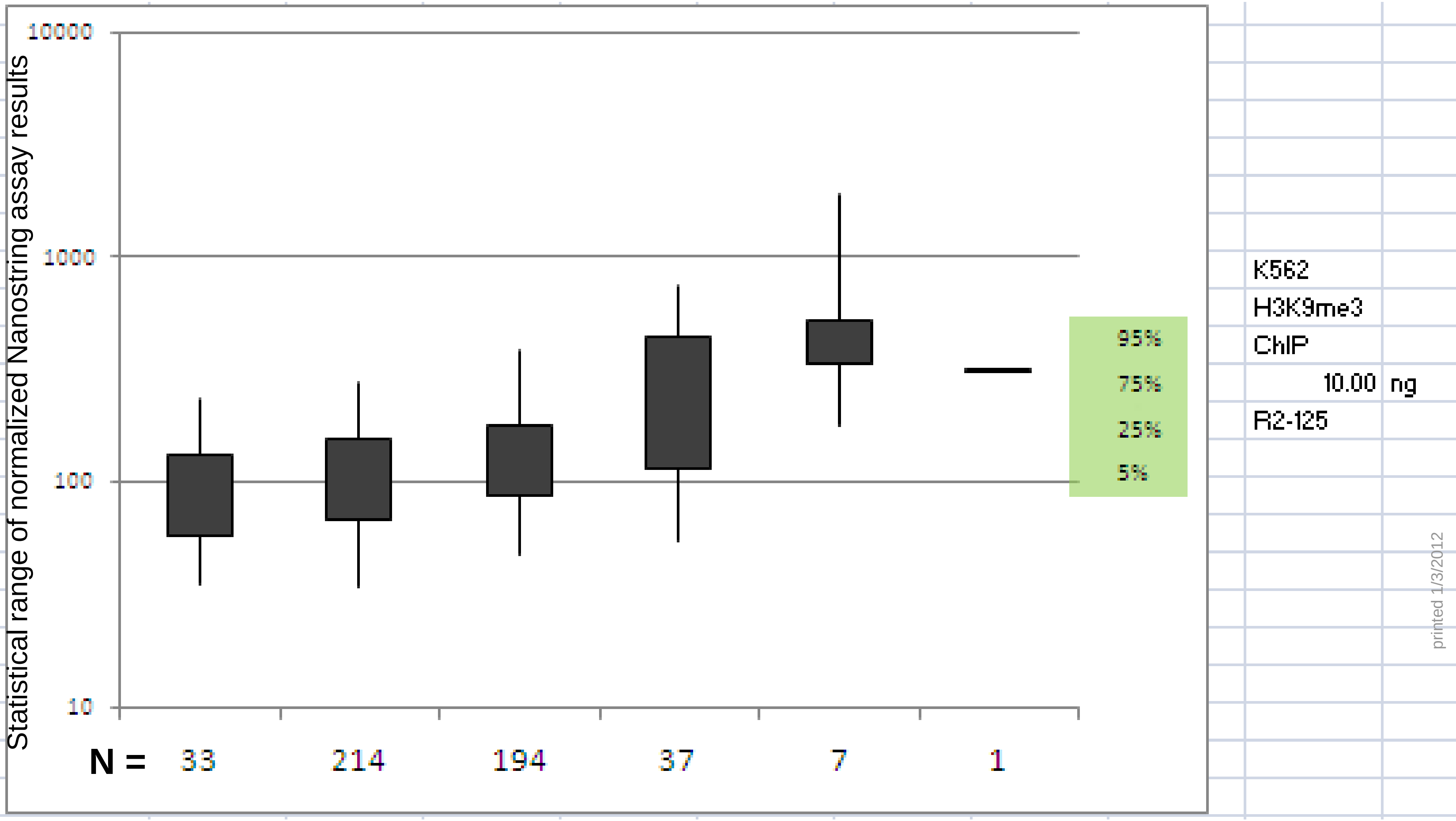
**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2  2 – 5  5 – 10  10 – 20  20 – 40  > 40**

# K562 H3K79me2 library DNA



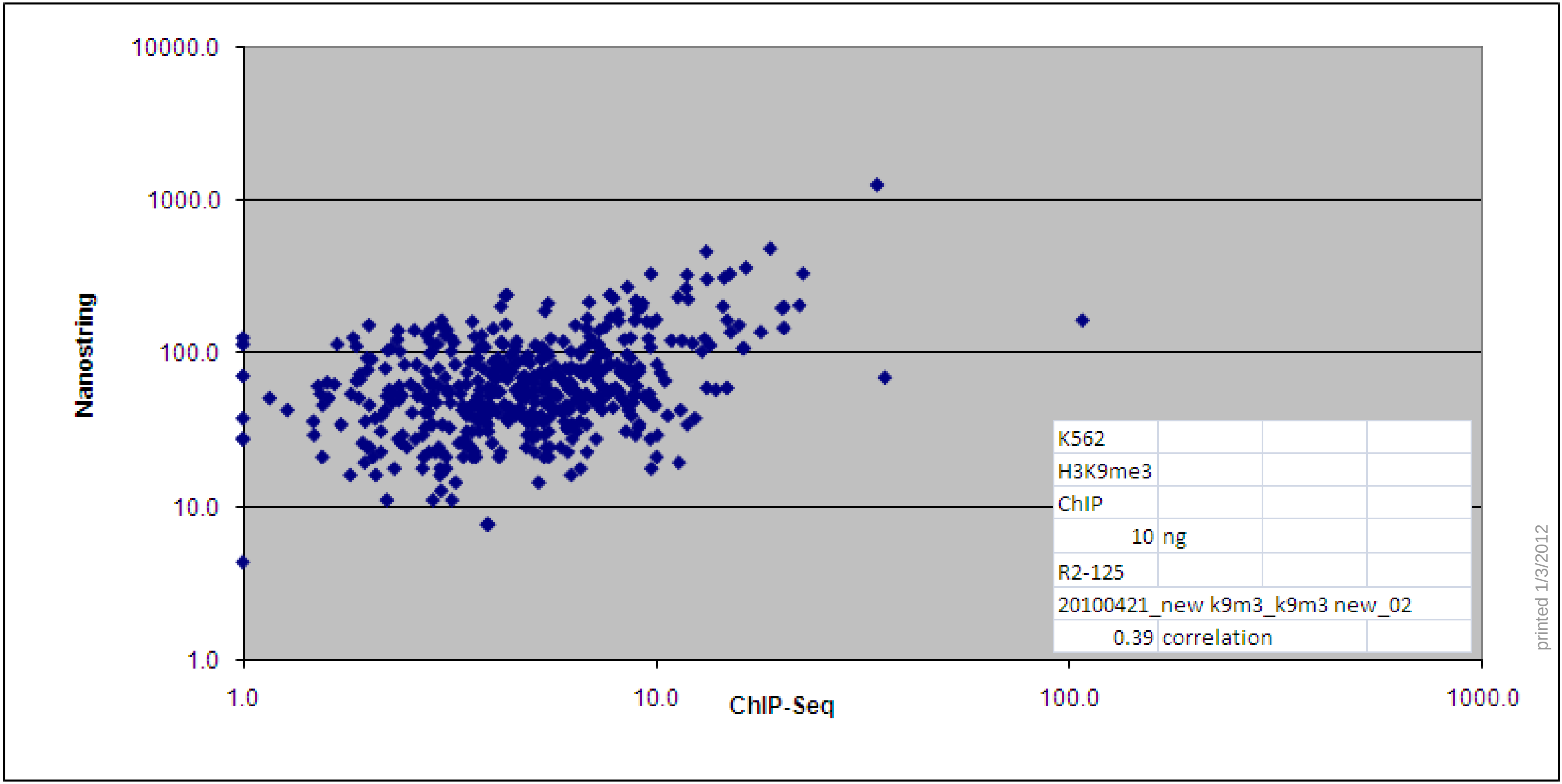
printed 1/3/2012

# K562 H3K9me3 ChIP



**Fold Enrichment from ChIP-Seq (normalized intensity)**
**< 2 2 – 5 5 – 10 10 – 20 20 – 40 > 40**

# K562 H3K9me3 ChIP



| | | | |
|---|---|---|---|
| K562 | | | |
| H3K9me3 | | | |
| ChIP | | | |
| 10 | ng | | |
| R2-125 | | | |
| 20100421_new k9m3_k9m3 new_02 | | | |
| 0.39 | correlation | | |

printed 1/3/2012