Fully automated ChIP on two Histones using Diagenode's iDeal ChIP-seq kit for Histones in Inorevia's Magelia® platform

Sebastian Aguilar Pierlé[1], Magda Granata[1], Nathalie Sondermann[2], Irina Panteleeva[2], Camille Soucies[1], Sol Schvartzman[2], Chloé Galbert[1], David Gosselin[1], Julien Autebert[1], Agnieszka Zelisko-Schmidt[2] and Amel Bendali[1]

1 Inorevia, 29 rue du Faubourg St Jacques, 75014 Paris, France – www.inorevia.com / contact@inorevia.com
2 Diagenode, Rue Bois Saint-Jean, 3, 4102 Seraing (Ougrée), Belgium

INTRODUCTION

ChIP-seq is nowadays the method of choice for identification of protein-DNA interactions at a whole genome scale. Despite its widespread use, this challenging application has significant limitations to hurdle, including the requirement for relatively large amounts of input material, reproducibility, recovery, and increased signal-to-noise ratio. Miniaturization paired with automation represent a tangible solution to tackle these challenges.

Magelia®, Inorevia's multi-OMICs platform is fully compatible for automation of Diagenode's iDeal ChIP-seq Kit for Histones (Diagenode, Cat. No. C01010051). Its precise handling of a variety of magnetic beads paired with starting material reduction are adapted for treating precious samples. Reduced reaction volumes improve kinetics for an efficient IP. This in turn enables overall time reduction for long incubations in the platform. Here, we show ChIP performance using H3K4me3 and H3K9me3 antibodies (Diagenode, Cat. Nos. C15410193 & C15200152) in Magelia® with a 40x initial material reduction.

KEY BENEFITS OF Magelia® TECHNOLOGY

- **Full automation of complex workflows**
- **High efficiency on low quantities**
- **Time savings on hands-on time & turn around time**
- **Automated reagent aliquoting**
- **Compact benchtop multi-OMICs platform**
- **Unparalleled precision: no cross-contamination, no evaporation**

METHODS

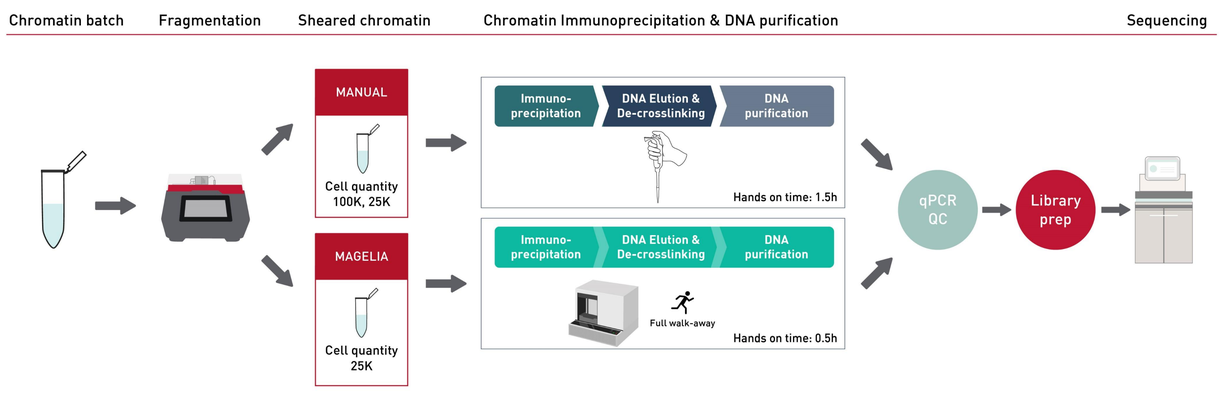

FIGURE 1 : Magelia® CHIP WORKFLOW

ChIP experiments were performed on K562 cell chromatin. The chromatin batch was prepared according to Diagenode's iDeal ChIP-seq kit for Histones recommendations. Total material equivalent to 25 000 cells was processed for ChIP in parallel manually and in the Magelia® respectively, following the manufacturer's instructions (Figure 1). In addition, 100 000 cells were processed manually as a control. Samples were treated in duplicate for H3K4me3 and triplicate for H3K9me3.

Immunoprecipitated DNA was then analyzed through qPCR using EIF4A2 and TSH2B primers provided by Diagenode to confirm the enrichment of the samples. Samples were then prepared for sequencing using the MicroPlex Library Preparation Kit v3 from Diagenode (Cat. No. C05010001). Samples were then sequenced using a NovaSeq 6000 S1 cartridge, PE 2x50 bp.

FASTQ files were evaluated using the FastQC tool for quality control inspection. The reads were then trimmed and mapped to the Homo sapiens (human) genome assembly GRCh37 (hg19) version.

Although most analyses were performed for both marks, for ease of visualization, one mark is shown at a time for specific analyses. These are detailed in the following paragraph and the results.

Relevant parameters for QC as published by the Encyclopedia of DNA Elements (ENCODE) consortium, were applied. Reproducibility was determined by verifying that 80% of the top 40% peaks overlapped between replicates. Peaks were ranked in function of statistical significance. Peak density per chromosome was determined for H3K4me3. To confirm previously known genes that interact with H3K9me3 on cell line K562, we referred to Blahnik et al.'s 2011 study. In this publication, it was shown that gene ZNF555 is bound by H3K9me3 in the aforementioned cell line, conversely it was demonstrated that gene GAPDH was not. This enabled us to confirm these findings using these genes as positive and negative controls, respectively.

RESULTS

ChIP recovery

Figure 2 shows average recovery for H3K4me3 and H3K9me3 replicates. Overall, recovery is comparable for both marks/conditions. Magelia® treated samples showed lower variability and higher recovery. No significant differences were seen between conditions, except for H3K9me3 recovery, with Magelia® treated samples showing higher values (p=0.027).

iDeal ChIP-seq kit for Histones in Magelia® produces comparable sequencing data for H3K9me3

Quality sequencing data was obtained for all samples as shown on Table 1 (H3K4me3 not shown). Samples across conditions showed a minimum of 98,05% of reads with a PHRED score above 30. All samples showed >99% of reads mapped. Generally, moderate percentages of duplication were detected across samples/conditions, except for the Magelia input sample from replicate 2 (>60%).

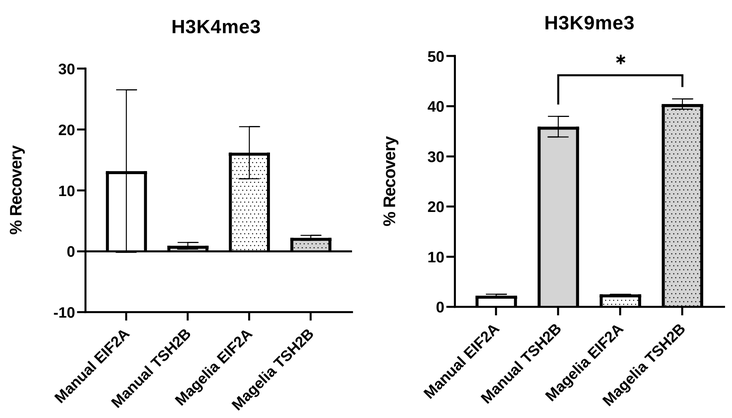

FIGURE 2: AVERAGE RECOVERY FOR H3K4ME3 AND H3K9ME3 REPLICATES.
Checkered bars represent Magelia® treated samples.

All Magelia treated samples reached 45 million uniquely mapped reads, aligning with ENCODE recommendations. This was not the case for manually treated replicate 1. This highlights the compatibility of Magelia® with the iDeal ChIP-seq kit for Histones.

Replicate	Sample	Total reads	% Reads Phred score> 30	% reads mapped	% duplicates	UMR
1	INPUT Man	78 735 621	99,02	99,13	43,95	44 131 315
	IP K9 100K Man	55 673 402	98,05	99,5	33,33	37 117 457
	IP K9 25K Man	40 893 151	98,87	99,27	43,46	23 120 987
	INPUT Mag	99 271 998	98,93	99,84	15,31	84 073 455
	IP K9 25K Mag	83 540 122	98,69	99,53	44,51	46 356 413
2	INPUT Man	128 661 806	98,98	99,51	47,87	67 071 399
	IP K9 100K Man	105 598 680	98,69	99,35	42,75	60 455 244
	IP K9 25K Man	97 574 546	98,78	99,89	24,63	73 541 935
	INPUT Mag	113 496 264	98,95	99,01	60,01	45 387 155
	IP K9 25K Mag	118 279 240	98,92	99,84	18,84	95 995 431
3	INPUT Man	80 092 826	99,1	99,19	28,82	57 010 073
	IP K9 100K Man	74 992 228	98,86	99,27	39,42	45 430 291
	IP K9 25K Man	86 598 128	99,01	99,15	30,38	60 289 616
	INPUT Mag	98 418 320	98,87	99,91	25,50	73 321 648
	IP K9 25K Mag	88 370 978	98,91	99,88	20,20	70 520 040

TABLE 1: SEQUENCING METRICS FOR H3K9me3 REPLICTAES

Peak correspondence for H3K4me3

Figure 3 shows assigned peaks across chromosome 1 of the human genome represented by blue bars as an example. The percentage correspondence for targets across the whole analysis for all samples/replicates is reported.
As shown on this figure, replicates 1 and 2 of the 25K tests were found to have above 80% correspondence across the board, highlighting the technical reproducibility of this experiment. This was true for both marks (H3K9me3 not shown).

Positive and negative control genes for H3K9me3

Figure 4 shows the peak signal for gene ZNF555, an example of a gene known to be bound by H3K9me3, and GAPDH, a gene used as a negative control (Blahnik et al. 2011). The shade gradient represents the signal (assigned peaks) detected for a given gene, with the darkest coloring being the lowest value and the lightest the highest (0 to >2800). Each column represents a condition and each row a gene. Known positive (ZNF555) and negative control (GAPDH) genes are highlighted in a grey square. Signal intensity confirms the expected results for both positive and negative control genes, highlighting the concordance of the obtained data across conditions/replicates with previously published work on this cell line.

Peak density across Chromosomes for H3K4me3

Zhao and colleagues investigated in 2007 the genome wide mapping of H3K4me3.

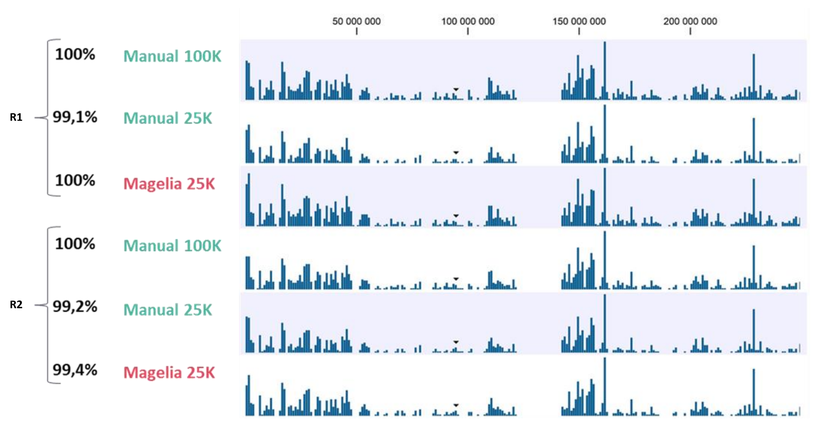

FIGURE 3: PERCENTAGE CORRESPONDENCE OF TOP 40% PEAKS ACROSS REPLICATES FOR H3K4me3. Full-length chromosome 1 is shown as an example. The coordinates are indicated on top. Blue bars represent areas with peaks assigned, peak height is proportional to peak density. R1 and R2 indicate replicates 1 and 2.

One of their findings was that gene density on a chromosome level was correlated with H3K4me3 cluster density. Logically, chromosomes 18 and 19, which have the lowest and highest gene densities respectively, also displayed the lowest and highest H3K4me3 cluster density.
We performed the same analysis on our data set. As seen on Figure 5, the same trends were found across samples and replicates. Although 100k samples are not shown, the same profiles were identified. **<u>This further reinforces the technical concordance of our ChIP-seq dataset as a relation between gene and peak density was identified genome-wide.</u>**

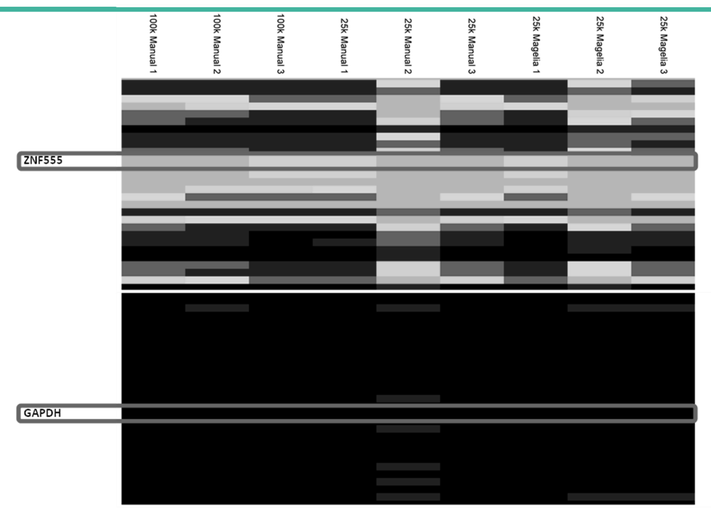

FIGURE 4: SIGNAL FOR SELECTED POSITIVE AND NEGATIVE CONTROL GENES FOR H3K9me3.

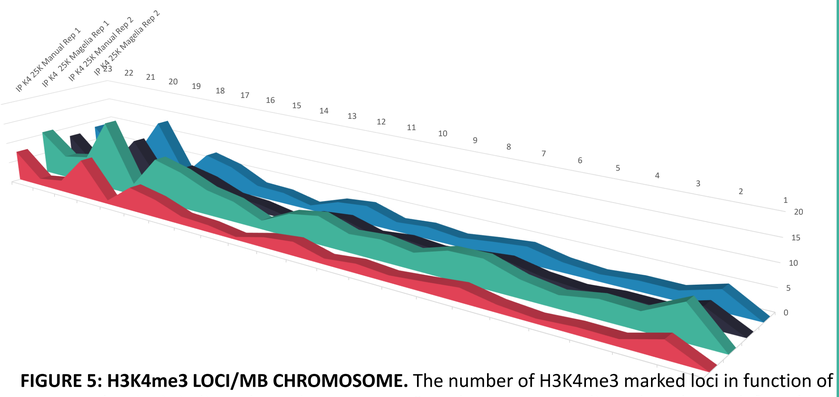

FIGURE 5: H3K4me3 LOCI/MB CHROMOSOME. The number of H3K4me3 marked loci in function of gene density is indicated per chromosome. Sample names are indicated on the top left and chromosome numbers (in reverse order) on the top x axis. The Y axis on the right indicates loci/mb.

CONCLUSIONS

The iDeal ChIP-seq kit for Histones paired with Magelia®, embodies an effective, and automated method to tackle a complex application:

- Magelia® treated samples performed comparably to (and sometimes better than) manually treated samples.
- This technical validation allowed to confirm previously published findings on the behavior of the studied marks.

The iDeal ChIP-seq kit for Histones transposed to Magelia® is now part of our ever-growing list of applications, allowing researchers and clinicians to harness its robustness, while reducing starting material quantities.

PERSPECTIVES

We are continuously developing a variety of miniaturized applications optimized for low input samples:

- RNA-seq lib prep with rRNA depletion/mRNA capture, challenging the limits of material required and securing more material with the same amount of starting RNA, while reducing amplification cycle number.
- WGS PCR-free library prep with and without enzymatic fragmentation, challenging the low limit of material required and reducing adapter dimer ratios.