Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells

© 2010 Nature America, Inc. All rights reserved.

To the Editor:

In a recent study in *Nature Genetics*, Wen *et al.*[1] reported the intriguing observation that the human and mouse genomes are organized into large histone H3 lysine 9–dimethylated (H3K9me2) chromatin blocks in various cell types. One of the main conclusions drawn by the authors is that such large domains are minimally present to absent in undifferentiated embryonic stem (ES) cells but arise upon differentiation. We claim that this conclusion is not supported by the authors' data.

In Figure 3a of their report, Wen *et al.* showed a map of discretized H3K9me2 domains that were defined by applying a threshold method. This map suggested that H3K9me2 domains are rarer, and much smaller, in undifferentiated ES cells than in differentiated ES cells. However, plotting of the authors' original H3K9me2 signals shows that the H3K9me2 domains are in fact very similar between the two cell types (**Fig. 1a**,**b** and **Supplementary Fig. 1**). It is clear from visual inspection of the plots shown in **Figure 1** and **Supplementary Figure 1** that using a slightly lower cutoff threshold would have resulted in much greater concordance of domains between cell types. Hidden Markov model (HMM) algorithms are routinely used for detecting chromatin domains in binding data[2–4], primarily because they do not rely on user-defined parameters and are therefore more objective. Similar to the binary classification (LOCK or not LOCK) that Wen *et al.* applied, we applied a two-state HMM to the original datasets. We found that this standard HMM algorithm identifies domains in undifferentiated and differentiated ES cells with roughly equal coverage of the genome and with overall 84.2% agreement between the two cell types (**Fig. 1c**).

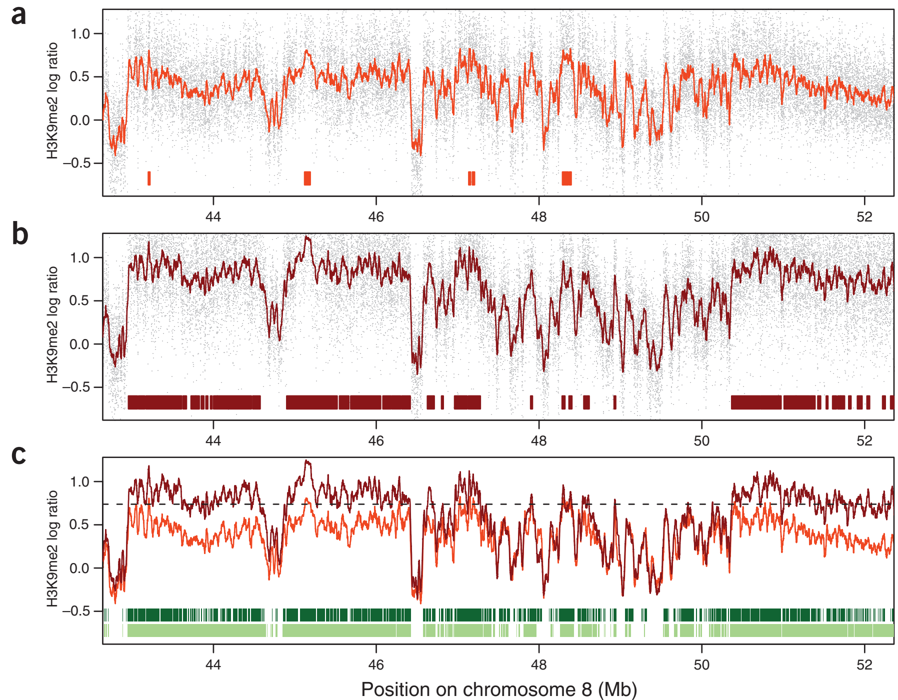

Figure 1 Reanalysis of H3K9me2 data in undifferentiated and differentiated ES cells. (**a**,**b**) Overview of the original normalized H3K9me2 data[1] (Gene Expression Omnibus accession code GSE13445) showing the same chromosome segment as displayed in Figure 3a of Wen *et al.*[1] in undifferentiated (**a**) and differentiated (**b**) ES cells. Gray dots represent data from individual array probes; bright and dark red curves show sliding window average (window size 93 probes). Rectangles indicate positions of discrete H3K9me2 domains ('LOCKs') as identified by Wen *et al.* (**c**) The same curves as in **a** and **b** are shown together with the approximate threshold value (dashed line) as applied by Wen *et al.* Light and dark green rectangles indicate discrete H3K9me2 domains that we identified using a two-state HMM algorithm in undifferentiated and differentiated ES cells, respectively. Emission probabilities were modeled by a *t* distribution. For each experiment, all the parameters were fitted to the data by the Baum-Welch algorithm, assuming the same standard deviation and degrees of freedom for both states. The genome coverage of H3K9me2 domains, as defined by the HMM algorithm, is 60.2% and 60.5% in undifferentiated and differentiated ES cells, respectively.

Such a strong discrepancy in results between the two methods calls for closer examination of the data. Inspection of the locus shown in Figure 3a of Wen *et al.* and many other randomly selected regions (**Fig. 1a**,**b** and **Supplementary Fig. 1**) shows that the qualitative differences they reported rely on a relatively subtle difference in dynamic range between their two experiments. However, the authors did not produce controls showing that this difference in dynamic range is reproducible and not purely technical. Particularly in the absence of replicates, microarray experiments are prone to artifacts. Variations in sample labeling or hybridization conditions, for instance, can easily account for the 0.2–0.4 $\log_2$ unit differences they observe (**Supplementary Fig. 1**). Moreover, the authors did not rule out systematic experimental errors related to the chromatin immunoprecipitation procedure, such as differences in epitope accessibility or locus-dependent variation in the efficacy of chromatin extraction between cell types. The latter is commonly controlled for by conducting a parallel genome-wide mapping experiment with histone H3 antibody.

In summary, we believe it is premature to conclude that there is a fundamental difference between H3K9me2 domains in undifferentiated ES cells and differentiated cells. Rather, based on the data presented by Wen *et al.*, we conclude that the domains appear to exist in both cell types with roughly equal coverage of the genome. The reproducibility and biological importance of the differences requires further investigation.

Guillaume J Filion & Bas van Steensel

Division of Gene Regulation, Netherlands Cancer Institute, Amsterdam, The Netherlands. Correspondence should be addressed to B.v.S. (b.v.steensel@nki.nl).

Note: Supplementary information is available on the Nature Genetics website.

AUTHOR CONTRIBUTIONS

G.J.F. and B.v.S. analyzed data and wrote the manuscript.

1. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A. & Feinberg, A.P. *Nat. Genet.* **41**, 246–250 (2009).
2. Birney, E. *et al. Nature* **447**, 799–816 (2007).
3. Koch, C.M. *et al. Genome Res.* **17**, 691–707 (2007).
4. Humburg, P., Bulger, D. & Stone, G. *BMC Bioinformatics* **9**, 343 (2008).