Unravelling the role of NANOG in the regulation of priming gene expression

Raquel Calçada[1,2,*]

Thesis to obtain the Master of Science Degree in Biomedical Engineering – November, 2016

Supervisors: Prof. Dr. Domingos Henrique[2] and Prof. Dr. Margarida Diogo[1]

ABSTRACT: The transcription factor NANOG exhibits a heterogeneous expression in pluripotent mouse embryonic stem (mES) cells, both at mRNA and protein levels. This variety appears to arise from stochastic fluctuations in NANOG expression in individual mES cells, creating windows of opportunity to explore pluripotency. Low-NANOG cells, in a "lineage-primed" state, are more susceptible to commit to differentiation and express higher levels of lineage-affiliated genes, than high-NANOG cells.

However, it is not clear how NANOG controls the exit of mES cells from pluripotency. Recent studies revealed that NANOG might interact with polycomb repressive complex 2 (PRC2) and ten-eleven translocation 1 (TET1), responsible for H3K27 trimethylation and DNA demethylation, respectively. Moreover, genes upregulated in low-NANOG cells, denoted as priming genes, which include lineage-affiliated genes, are enriched for binding signatures in PRC2. Thus, it was hypothesized that NANOG represses priming gene expression by regulating PRC2 and TET1, but how?

To unravel these mechanisms, the cell lines E14tg2a and Nd (with a Nanog:VNP reporter) were cultured in pluripotent "Serum/LIF" conditions. mES cells were incubated with GSK343, an inhibitor of PRC2 activity, and Ascorbic Acid (AA), a promoter of TET1. Afterwards, the effects on the expression of priming genes were evaluated using single-molecule RNA FISH. In accordance with preliminary data, now confirmed, GSK343 increases lineage-affiliated gene expression in high-*Nanog* cells, whilst AA decreases. Thus, it is proposed that NANOG forms a complex with TET1, which by maintaining a hypomethylated state, contributes to PRC2 recruitment and, consequently, to the silencing of lineage-affiliated gene expression.

KEY WORDS: NANOG; embryonic stem cells; pluripotency; heterogeneity; PRC2; TET1.

I. INTRODUCTION

Mouse embryonic stem (mES) cells are derived from the inner cell mass (ICM) epiblast of the mouse embryo blastocyst. mES cells are characterized by two main hallmarks: the self-renewal divisions and pluripotency with the capacity to differentiate into all embryonic cell lineages of an embryo. These mES cells can be maintained in pluripotency under specific conditions, such as "Serum/Leukemia Inhibitory Factor" (LIF) medium. ES cells constitute a very promising resource for regenerative medicine as well as for drug screening assays.

Nevertheless, it is essential to unravel the molecular mechanisms underlying pluripotency and how cells exit into differentiation.

It is known that pluripotency is governed by a gene regulatory network composed of three core transcription factors: NANOG, OCT4 and SOX2. Together, they allow mES cells to test multiple lineage programmes before commitment, while ensuring that the pluripotency state is sustained. Contrarily to OCT4 and SOX2, which exhibit a homogeneous protein expression in ES cell cultures, as well as in the blastocyst's ICM, NANOG presents a heterogeneous expression (Figure 1), with approximately (56.2 ± 8.0)% of Nd cells expressing a Nanog:VNP reporter, when grown in "Serum/LIF" (Abranches *et al.* 2013).

[1]Instituto Superior Técnico, University of Lisbon, Portugal

[2] Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, Portugal

* Email: raquel.calcada@tecnico.ulisboa.pt

NANOG heterogeneity arises from stochastic fluctuations in NANOG expression in individual mES cells, as observed by time-lapse video and also corroborated by the mRNA heterogeneity detected by single-molecule RNA fluorescence *in situ* hybridization (smFISH) (Abranches *et al.* 2014).

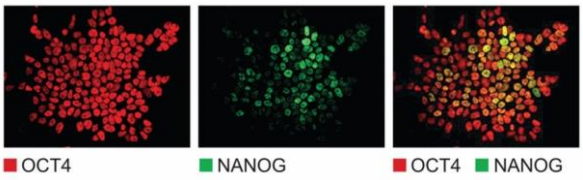


Figure 1 – NANOG heterogeneity. *Contrarily to OCT4 homogeneous expression, not all mES cells express NANOG. Adapted from Torres-Padilla & Chambers 2014.*

ES cells maintain a balance between self-renewal and differentiation, in other words, between pristine pluripotency and lineage specification. This equilibrium is influenced not only by environmental cues, supplied by the culture medium, but also by intrinsic transcription factors, such as NANOG (Figure 2) (Osorno & Chambers 2011).

The fact that *Nanog*$^{-/-}$ mES cells display an increased predisposition to differentiation (Chambers *et al.* 2007), led to the hypothesis that the low-NANOG cells are in a "lineage-primed" state, with a higher tendency to differentiate, less self-renewal and an increased expression of lineage-affiliated genes when compared to high-NANOG cells (Abranches *et al.* 2013).

"Lineage priming**"** is a process, within pluripotency, during which mES cells display an increased but reversible expression of lineage-affiliated genes, in a low-NANOG state, reflecting a higher predisposition to commit into a specific cell-fate. It is "the seeding of a particular fate on the way to commitment" (Martinez Arias *et al.* 2013).

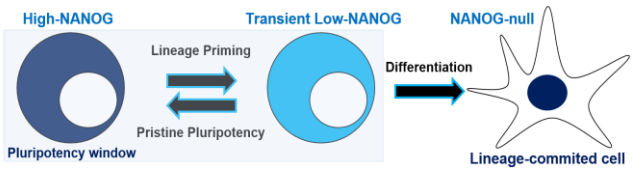


Figure 2 – Nanog as a "differentiation rheostat". *Nanog is constantly fluctuating between high and low states. Cells which overexpress Nanog are skewed towards self-renewal whist Nanog-null cells tend to differentiate. A transient Nanog negative cell might be in a "lineage-primed" state, more prone to commitment, although its decision may be reverted by the re-expression of Nanog. Adapted from Chambers* et al. *2007.*

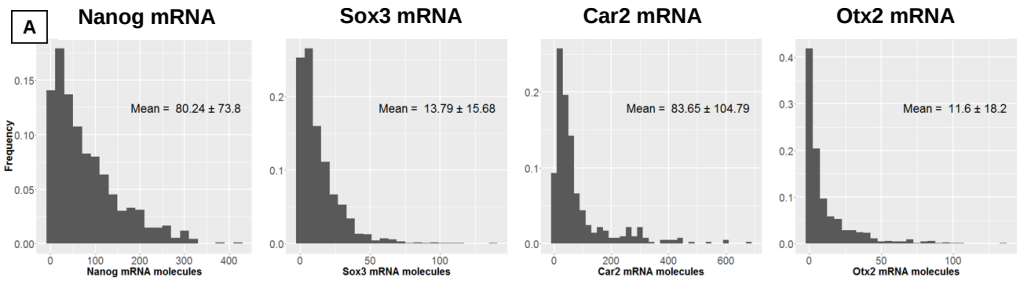
Previous work from DHenrique Lab found, by RNA-sequencing, a wider class of genes upregulated in low-Nanog:VNP cells, denoted as priming genes, which comprise not only lineage-affiliated genes, which are specific for a lineage programme, but also sporadic genes with unknown role in development.

But how does NANOG regulate priming gene expression? It is known that NANOG interacts with PRC2 (Gagliardi *et al.* 2013) and TET1 (Costa *et al.* 2013), involved in the trimethylation of histone 3 at lysine 27 (H3K27me3) and in DNA demethylation through the conversion of 5mC into 5hmC, respectively (Aloia *et al.* 2013; Kohli & Zhang 2013).

Interestingly, priming genes show enriched binding signatures for PRC2 components (unpublished results from DHenrique Lab). Thus, it was proposed that the PRC2 complex might be involved in the repression of priming gene expression in a high-NANOG state through the trimethylation of H3K27 on their promoters. GSK343 is a specific inhibitor of EZH2, the catalytic subunit of PRC2, responsible for blocking the *de novo* trimethylation of H3K27, which is predicted to occur from low- to high-NANOG state.

Furthermore, it was hypothesized that TET1 forms a complex with NANOG (Costa *et al.* 2013) which indirectly, by maintaining a hypomethylated state at CpG-rich promoters, contributes to PRC2 recruitment to these binding sites, leading to the silencing of priming gene expression (Wu *et al.* 2011; Sui *et al.* 2012). Ascorbic acid (AA), also known as vitamin C, promotes TET1 activity, leading to a rapid and global increase of 5hmC levels (Blaschke *et al.* 2014; Yin *et al.* 2013).

The aim of this project is to verify this model of NANOG regulation of priming gene expression, using the small molecules described (GSK343 and AA), that interfere with PRC2 and TET1 activity, respectively, as well as taking advantage of the single-cell analysis of smFISH.



B

Gene	Mean	SD	Min	Max	FF	N
Nanog	80.24	73.80	0.00	426.00	67.88	1088
Sox3	13.79	15.68	0.00	141.00	17.82	1088
Car2	83.65	104.79	0.00	686.00	131.27	408
Otx2	11.60	18.20	0.00	137.00	28.54	680

Figure 3 – Analysis of gene expression in E14 mES cells cultured in "Serum/LIF". ***(A)*** *Histograms of the distribution of mRNA transcripts/cell for Nanog, Sox3, Car2 and Otx2.* ***(B)*** *Statistical measurements for Nanog, Sox3, Car2 and Otx2. The parameters shown are the mean, standard deviation (SD), minimum (Min), maximum (Max), Fano factor (FF) and the number of cells analysed by smFISH.*

II. RESULTS AND DISCUSSION

A. Nanog and Lineage Priming

In order to assess NANOG protein expression through flow cytometry analysis, it was used the Nd mES cell line, which contains a dynamic fluorescent Nanog:VNP reporter and was generated from E14 cells (Abranches *et al.* 2013).

As it was reported before for Nd mES cells cultured in "Serum/LIF", there are approximately 50% of Nanog:VNP positive cells (data not shown). This protein heterogeneity is linked to stochastic NANOG fluctuations in individual cells and caused by a striking *Nanog* mRNA variability.

In order to analyse the correlation between high priming gene expression and *Nanog* in individual mES cells, smFISH was performed. The smFISH analysis allows to obtain the number of mRNA molecules of each gene per individual cell, within a large population (Raj *et al.* 2006). For each gene it is possible to obtain histograms of the mRNA distribution relatively to the cell frequency (Figure 3.A) and also a statistical analysis, with mean, standard deviation (SD), minimum (Min), maximum (Max), Fano factor (FF) and number of cells (N) (Figure 3.B).

Amongst the priming genes identified by RNA-sequencing are *Sox3* and *Car2*. *Sox3* is a lineage-affiliated gene involved in neural differentiation whilst *Car2* is a sporadic gene that encodes for a carbonic anhydrase, with unknown role in development and sporadic expression. On the other hand, *Otx2* is a marker of the transition between ES cells to epiblast stem cells, which establishes a parallel with pre- to postimplantation embryo (Acampora *et al.* 2013).

All the depicted genes present a long-tailed mRNA distribution (Figure 3.A) and varying mRNA levels, for instance *Nanog* ranges from 0 to 426 mRNAs/cell, with an average of 80 mRNAs/cell (Figure 3.B).

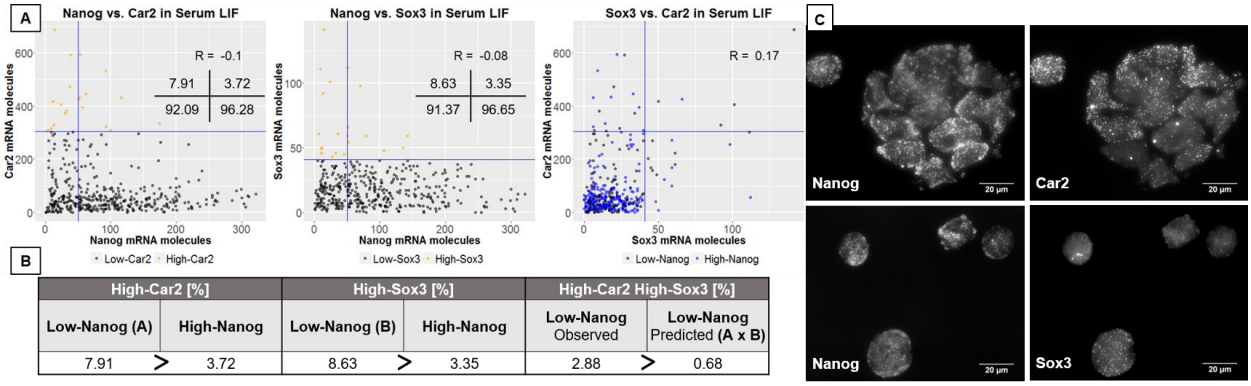
These mRNA distributions indicate that the majority of cells has none or very few mRNA transcripts of the respective gene (inactive transcription), whilst there are few cells expressing variable and high levels (active transcription) (Figure 3.A).

Furthermore, genes also present a Fano factor higher than 1, typical of non-Poisson models (Figure 3.B). These results suggest a bursty or pulsatile transcription.

The high or low state of each gene was defined by a specific threshold. For *Nanog* the threshold is 50 mRNAs/cell and for the other genes the top-5% expressing cells are selected for the high-state, depicted in orange (Figure 4.A).

The majority of cells exhibits lower levels of both priming genes: 86.33% of low-*Nanog* cells and 93.31% of high-*Nanog* cells (data not shown).

Nevertheless, within the minority, it is observed that, when compared to high-*Nanog* cells, there is a higher percentage of low-*Nanog* cells expressing increased levels of *Car2* (7.91% > 3.72%) and *Sox3* (8.63% > 3.35%) transcripts



High-Car2 [%]		High-Sox3 [%]		High-Car2 High-Sox3 [%]	
Low-Nanog (A)	High-Nanog	Low-Nanog (B)	High-Nanog	Low-Nanog Observed	Low-Nanog Predicted (A x B)
7.91 >	3.72	8.63 >	3.35	2.88 >	0.68

Figure 4 – Correlation analysis of priming gene expression in E14 mES cells cultured in "Serum/LIF". (A) *Correlation scatterplots for Car2, Nanog and Sox3. R is the Spearman correlation coefficient and the numbers correspond to the percentage of cells expressing higher or lower levels of priming gene within the respective Nanog subpopulation: low-Nanog cells on the left and high-Nanog cells on the right column.* ***(B)*** *Percentage of cells expressing higher levels of priming genes (Car2 and Sox3) within the low and high-Nanog subpopulations, suggesting the occurrence of "lineage priming" for Sox3. The last two columns are for low-Nanog cells expressing high levels of both priming genes simultaneously, comparing the observed to the expected results if they were independent events.* ***(C)*** *Nanog mRNA expression is depicted on the left and priming gene mRNA expression on the right. Car2 is analysed in the first row and Sox3 in the second row. Scale bar = 20 µm.*

(Figure 4.B). This result confirms that both priming genes are upregulated in a low-*Nanog* state (Figure 4.C). Besides, in the case of *Sox3*, it suggests the existence of "lineage priming" in a low-*Nanog* state.

Furthermore, there is a poor correlation between *Car2* and *Sox3* mRNA molecules, as it can be inferred by the very weak Spearman correlation coefficient (R = 0.17) (Figure 4.A). This implies that the expression of one gene is not a good indicator of the other, which might indicate that there is no common upstream regulatory pathway for both genes.

In order to test if high-*Sox3* and high-*Car2* expression in low-*Nanog* cells are independent events, the mathematical formula (1) was applied, in which P(A) is the probability of expressing higher levels of *Car2* in low-*Nanog* cells, P(B) of *Sox3* and P(A ∩ B) of both:

$$P(A) \times P(B) = P(A \cap B) \qquad (1)$$

In fact, 7.91% x 8.63% = 0.68%, which is lower than the observed 2.88%, concluding that these events are not independent (Figure 4.B). One hypothesis is that the observed dependency might be due to neighbouring chromatin effects.

By contrast, *Otx2* is not upregulated in low-*Nanog* cells, but almost identically distributed between *Nanog* states: 5.17% of low-*Nanog* cells and 4.82% of high-*Nanog* cells express high levels of *Otx2* transcripts, concluding that *Otx2* cannot be a priming gene (data not shown).

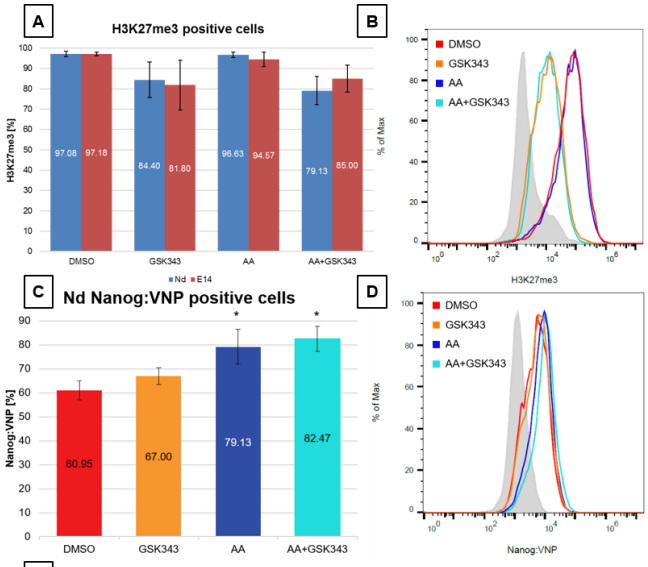
Moreover, when the independent event's mathematical formula (1) was applied, the predicted value was similar to the observed (data not shown), supporting the fact that the expression of *Otx2* and *Sox3* in low-*Nanog* cells are completely independent events, contrarily to what was detected for *Car2* and *Sox3*.

In sum, "lineage priming" is a rare event that occurs in a small fraction of low-*Nanog* cells that express high levels of lineage-affiliated genes (such as *Sox3*). "Lineage priming" in a transient low-*Nanog* state is essential since it not only allows mES cells to test lineage programmes (such as neural) before definitive commitment but also provides the opportunity of *Nanog* re-expression, thereby maintaining a pool of pristine pluripotent cells.

B. Nanog Regulation of Priming Gene Expression through PRC2 and TET1

In order to test the proposed model, mES cells were incubated during 48 hours with GSK343 and/or AA, always compared to the control DMSO (GSK343 solvent).

The first observation is that exposure of



E

	High-Car2 [%]		High-Sox3 [%]	
Condition	**Low-Nanog**	**High-Nanog**	**Low-Nanog**	**High-Nanog**
DMSO	7.32	3.63	7.64	3.07
GSK343	7.79	4.00	7.38	6.77
AA	19.48	10.22	7.79	1.78
AA+GSK343	11.20	9.62	7.20	5.13

Figure 5 – Analysis of the effects of GSK343 and/or AA. (A)*. Percentage of H3K27me3 positive cells obtained by IC-FC for E14 and Nd mES cells, treated with GSK343 and/or AA, and DMSO in parallel, for 48 hours.* ***(B)*** *Representative flow cytometry profile for Nd mES cells in different conditions, in which there is a shift to the left in cells treated with GSK343 (alone or combined). Samples only marked with secondary antibodies were used as negative fluorescent controls (depicted in filled grey).* ***(C)*** *Percentage of Nanog:VNP positive cells in Nd cells tested in different conditions. Statistically significant difference with p-value ≤ 0.05 (t test) is denoted with (*). On (A) and (C) mean values were depicted for each condition and error bars were calculated based on standard deviation from three experiments (n=3).* ***(D)*** *Representative flow cytometry profile in which there is an increase on Nanog:VNP upon AA exposure, visible by the shift to the right on cells treated with AA (alone or combined). E14 cells were used as a control (depicted in filled grey).* ***(E)*** *Percentages of low and high-Nanog cells that express high levels of Car2 (on the left) and Sox3 (on the right) transcripts for different conditions.*

GSK343 alone or combined with AA shows a reduction trend on the percentage of H3K27me3 expressing cells, when compared to DMSO, in both Nd and E14 cell lines (Figure 5.A), as it can be clearly observed by the shift to the left on the flow cytometry profile (Figure 5.B). This result, obtained by intracellular stainning - flow cytometry (IC-FC), confirms the efficient blocking of the *de novo* trimethylation of H3K27, performed by GSK343. The reduction is not more drastic because GSK343 only affects the *de novo* trimethylation and does not remove the histone repressive mark H3K27me3 that is already deposited in cells.

Moreover, the percentage of Nanog:VNP positive cells was assessed after 48 hours of GSK343 and/or AA exposure in Nd mES cells cultured in “Serum/LIF”. GSK343-treated cells have no significant changes on the Nanog:VNP percentages compared to the control, just a slight increase (60.95% for DMSO and 67% for GSK343) (Figure 5.C).

Nevertheless, the exposure of AA alone or combined with GSK343 increases the percentage of Nanog:VNP positive cells, when compared to the control (79.13% for AA and 82.47% for AA+GSK343) (Figure 5.C). The AA effect is also observed by the slight shift to the right on the flow cytometry profile (Figure 5.D). This result is line with recent studies which state that AA promotes TET1 activity, that in turn leads to an increased NANOG expression (Ito *et al*. 2010; Wu *et al*. 2011), as it will be explained in detail later.

These results are at a global scale, concerning the protein profiles, and serve as controls for the next step: the priming gene analysis in relation to *Nanog* mRNA expression by smFISH. Only a single-cell analysis will properly answer to our fundamental question of how mES cells undergo “lineage priming” through *Nanog* regulation of priming gene expression by interacting with PRC2 and TET1.

Therefore, mES cells were grown in “Serum/LIF” supplemented with GSK343 and/or AA, in parallel with DMSO, for 48 hours, then cells were fixed for smFISH and hybridised with the probe combination *Car2-Nanog-Sox3*.

mES cells treated with GSK343 show an increase on the percentage of high-*Nanog* cells expressing high levels of *Sox3* transcripts, when compared to DMSO (6.77% > 3.07%) (Figure 5.E). Nevertheless, there are no observable effects on the percentage of low-*Nanog* cells with increased expression of *Sox3* (7.38%), in

comparison to DMSO (7.64%). Interestingly, no effects are observed on *Car2* expression, relatively to DMSO (Figure 5.E).

These results suggest that GSK343, by inhibiting the *de novo* trimethylation of H3K27, prevents the repression of *Sox3*, a lineage-affiliated gene, expression in high-*Nanog* cells, which explains its increased expression. No effect was expected to occur in low-*Nanog* cells since GSK343 is only expected to act on the transition between low- to high-*Nanog* cells. The results suggest that this mechanism is only applicable to lineage-affiliated genes, such as *Sox3*, and not to sporadic genes, since *Car2* expression remains intact upon GSK343 treatment.

AA-treated mES cells show, by smFISH analysis, an overall increase on *Car2* mean expression, probably due to DNA demethylation mediated by TET1 on *Car2* promoter, leading to transcription activation. Moreover, there is also an increase of *Nanog* mean expression, which matches the observed increase on the percentage of Nanog:VNP positive cells upon AA exposure. There is a higher number of high-*Nanog* cells, when compared to DMSO, which reflects into a change on the histogram's shape of *Nanog* mRNA distribution, with the peak deviated to the right (data not shown).

mES cells treated with AA exhibit a slight decrease on the percentage of high-*Nanog* cells with high levels of *Sox3* transcripts, in comparison to DMSO (1.78% < 3.07%). Simultaneously, no effect is observed in low-*Nanog* cells with high *Sox3* expression (7.79%), compared to DMSO (7.64%). Contrarily, *Car2* expression globally increases in high-*Nanog* cells, which indicates that it is not regulated by the NANOG-TET1-PRC2 crosstalk and it is probably due to TET1-mediated DNA demethylation (Figure 5.E).

These results are in agreement with the proposed model for NANOG regulation combined with TET1, whose activity is promoted by AA. TET1 forms a complex with NANOG (Costa *et al.* 2013) and promotes DNA demethylation through the conversion of 5mC into 5hmC. These hypomethylated regions, which could be the promoters of lineage-affiliated genes, "attract" PRC2, recruiting it to its binding sites (Wu *et al.* 2011). Thus, EZH2 trimethylates H3K27, causing the repression of lineage-affiliated gene expression in high-NANOG cells. This model is in accordance with the fact that lineage-affiliated gene expression is upregulated in low-NANOG cells, where NANOG is absent and there is no formation of TET1-NANOG complex and consequently no 5hmC-mediated recruitment of PRC2. Thus, the model explains the existence of "lineage priming" in low-NANOG cells.

Finally, in AA+GSK343 condition, there is a similar result to the observed with GSK343. There is an increase on the percentage of high-*Nanog* cells which express high levels of *Sox3* transcripts, in comparison to DMSO (5.13% > 3.07%). Low-*Nanog* cells do not show any modifications and *Car2* expression globally increases due to TET1-mediated DNA demethylation (Figure 5.E).

This result corroborates the hypothesis that PRCS acts "downstream" of TET1, thus the effect of GSK343 prevails upon AA, it causes the inhibition of the recruited PRC2 and consequently prevents the repression of lineage-affiliated gene expression in high-*Nanog* cells.

C. Sorting into Low and High-Nanog:VNP Cells

To verify if histone *de novo* trimethylation of H3K27 occurs from low- to high-NANOG state, two sortings were performed. The sortings divided the bulk Nd population into low-Nanog:VNP (VNP_L) and high-Nanog:VNP (VNP_H) cells, which were replated in "Serum/LIF", with GSK343 and/or AA, in parallel to DMSO, for 48 hours.

Immediately after sorting, its efficiency was assessed by flow cytometry analysis of the percentage of Nanog:VNP positive cells on both subpopulations: 0.925% for VNP_L and 94.5% for VNP_H cells (Figure 6.A). Thus, the VNP_L and VNP_H subpopulations were considered pure.

Bright field images were taken to both FACS-sorted subpopulations 48 hours after GSK343 and/or AA exposure. In VNP_H cells there were tightly packed clusters with less flatten differentiated-like cells at the cluster's periphery or isolated between clusters, when compared to VNP_L cells. AA exposure in VNP_L cells also showed less differentiated-like cells in

comparison to other VNP_L cells cultured in different conditions (Figure 6.E).

The percentages of Nanog:VNP expressing cells obtained for DMSO, after 48 hours of exposure, were 14.45% for VNP_L and 56.85% for VNP_H (Figure 6.D), which are similar to those previously reported for "Serum/LIF" conditions. Moreover, VNP_H cells exhibited a faster restoration of NANOG heterogeneity when compared to VNP_L cells, as it was previously described (Abranches *et al.* 2013).

Relatively to VNP_L cells, GSK343 exposure during 48 hours has no notorious effect on the percentage of Nanog:VNP positive cells (16.45%), just a slight increase compared to DMSO (14.45%). However, when AA is applied, there is a significant increase on the percentage of Nanog:VNP positive cells (42.15% for AA and 52.40% for AA+GSK343) (Figure 6.B and Figure 6.D). Consequently, AA seems to promote the conversion from low- to high-NANOG state, thereby increasing the number of high-NANOG cells.

VNP_H cells treated with AA show a higher percentage of Nanog:VNP positive cells, when compared to DMSO (74.35% > 56.85%) (Figure

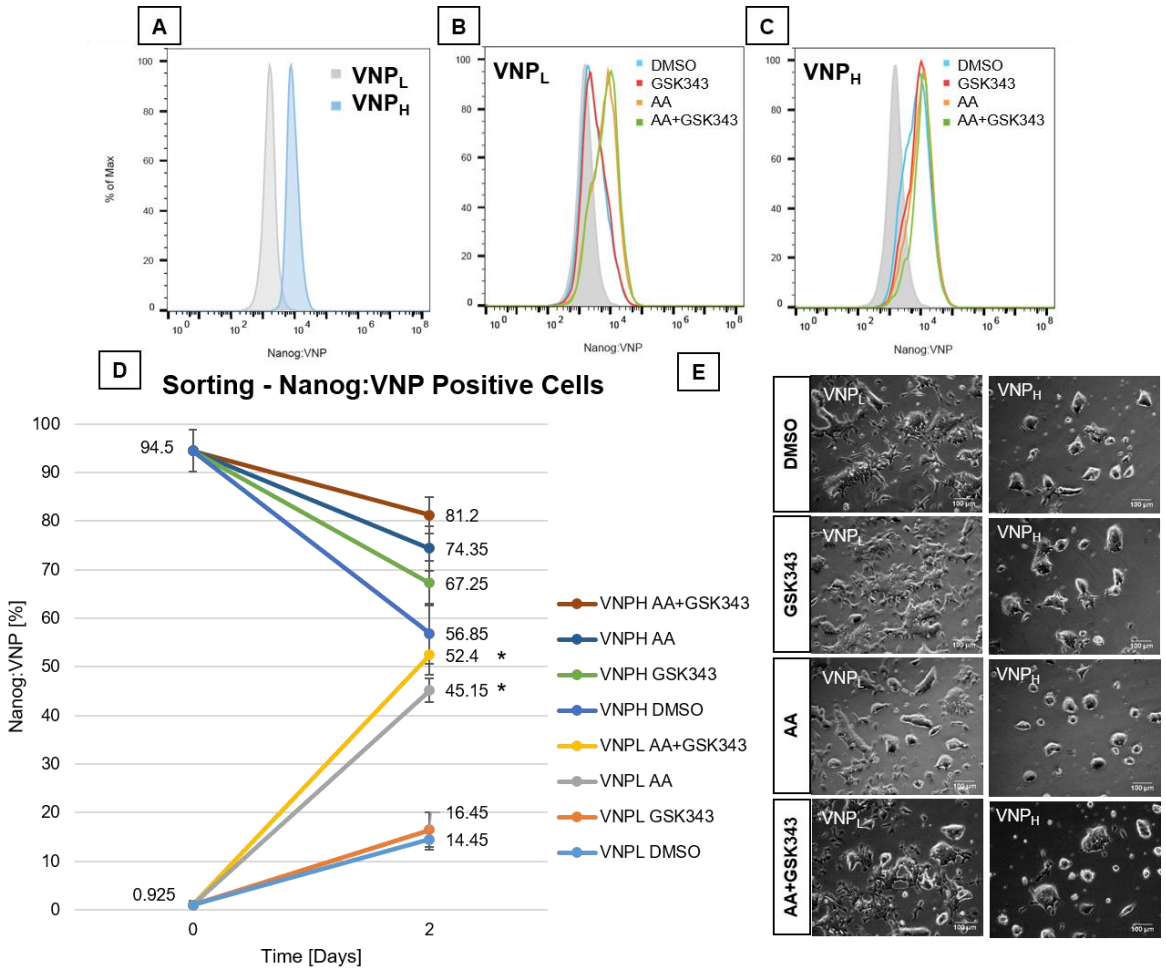


Figure 6 – Nanog expression in FACS-sorted Nd mES cells. (A) *Representative histogram of FACS-sorted Nd subpopulations VNP_L and VNP_H immediately after sorting.* ***(B)*** *Representative flow cytometry profile of VNP_L cells exposed to GSK343 and/or AA, in parallel to DMSO, for 48 hours.* ***(C)*** *VNP_H cells.* ***(D)*** *Temporal evolution of the percentage of Nanog:VNP positive cells after replating the FACS-sorted subpopulations in "Serum/LIF" supplemented with chemical modulators or DMSO, during 48 hours. Mean values were depicted for each condition and error bars were calculated based on standard deviation from two experiments (n=2). Statistically significant difference with p-value ≤ 0.05 (t test) observed between DMSO and respective VNP_L treated cells is denoted with (*).* ***(E)*** *Cell morphology of FACS-sorted subpopulations after 48 hours of exposure to chemical modulators or DMSO, before fixation for smFISH. Scale bar = 100 µm.*

6.C and Figure 6.D). Thus, AA also sustains the number of high-NANOG cells by blocking the transition from high- to low-NANOG.

In sum, AA increases the number of high-NANOG cells and the hypothesis to explain this results is based on two facts: 1) AA increases 5hmC levels through TET1 (Yin *et al.* 2013; Blaschke *et al.* 2014); 2) TET1 promotes the transcriptional of *Nanog* by maintaining a hypomethylated state, thus increasing NANOG expression (Ito *et al.* 2010; Wu *et al.* 2011).

Contrarily to the null effect upon VNP_L cells, GSK343 seems to maintain Nanog:VNP at higher percentages (67.25% for GSK343 and 81.2% for AA+GSK343) in VNP_H cells (Figure 6.C and Figure 6.D), when compared to DMSO.

The fact that GSK343 prevents the transition from high- to low-NANOG, and also the slight increase on Nanog:VNP described for bulk and VNP_L cells, are coincident with the fact that *Ezh2*-null iPS cells exhibit more NANOG expression (Villasante *et al.* 2011). Moreover, a recent study proved that GSK343 increases TET1 expression, which leads to higher NANOG levels, in a positive feedback loop (Neri *et al.* 2015).

These observations at the NANOG level serve as a reference for future smFISH experiments in VNP_L cells treated with GSK343 and/or AA. VNP_H cells will serve as a control, since priming gene expression is not expected to be affected in cells that transited from a high-NANOG state.

III. CONCLUSIONS

"Lineage priming" within pluripotency, is a very rare event restrict to a small fraction of low-*Nanog* cells. It consists on the upregulation of lineage-affiliated genes, such as *Sox3*, in low-*Nanog* cells, when compared to high-*Nanog* cells.

The smFISH results are in accordance with the hypothesis that GSK343 inhibits PRC2, blocking the *de novo* trimethylation of H3K27 and causing the consequent de-repression of lineage-affiliated gene promoters, justifying the increased

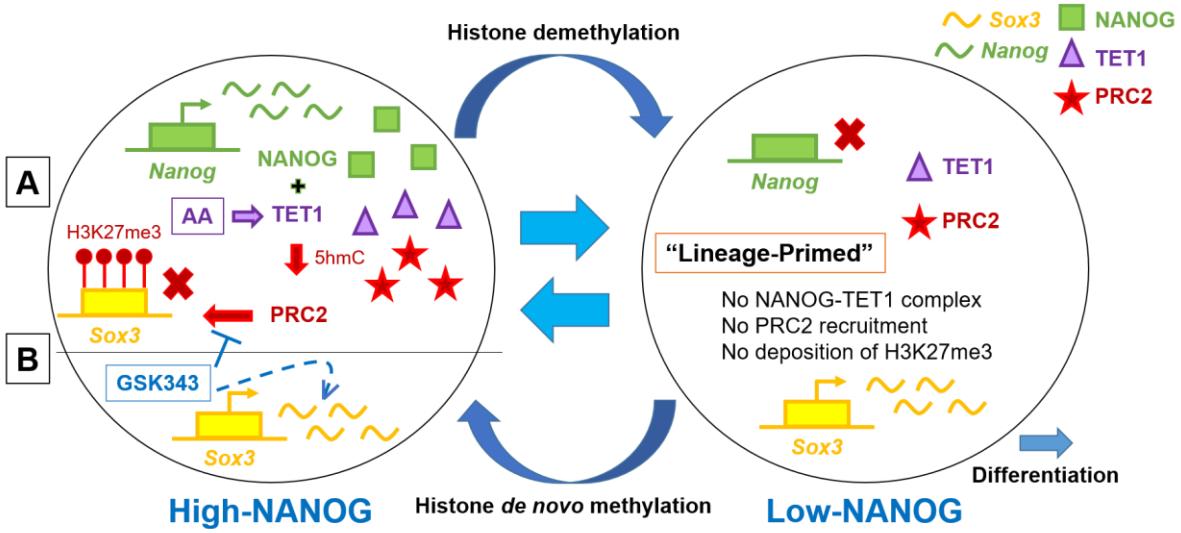


Figure 7 – Proposed model for NANOG regulation of lineage-affiliated gene expression and the effect of AA and GSK343 on priming. *NANOG presents stochastic fluctuations, transiting between high- and low-NANOG states. It is predicted that histone* de novo *methylation occurs from low- to high-NANOG state, due to PRC2 recruitment, "attracted" by hypomethylated regions promoted by the NANOG-TET1 complex. Consequently, PRC2 mediates the silencing of lineage-affiliated gene expression (represented by Sox3) in the high-NANOG state, through trimethylation of H3K27. When AA is added, it promotes TET1 action, leading to the decrease of Sox3 expression in high-NANOG cells. This represents what is expected to occur in a high-NANOG state (even without AA interference)* ***(A)****. GSK343, through EZH2 inhibition, leads to an increase of Sox3 expression in high-NANOG cells* ***(B)****. By contrast, in a low-NANOG state, NANOG is absent and there is no NANOG-TET1 complex, nor consequent PRC2 recruitment, which leads to an increase in Sox3 expression, related to "lineage priming".*

expression of *Sox3* in high-*Nanog* cells (Figure 7).

On the other hand, upon AA exposure, there is a decrease of *Sox3* expression in high-*Nanog* cells. This result is in line with the proposed model in which AA activates TET1, that forms a complex with NANOG (Costa *et al.* 2013) and this association will increase 5hmC levels. In turn, PRC2 is reported to be "attracted" to hypomethylated promoters (Wu *et al.* 2011), leading to H3K27 trimethylation and causing the repression of lineage-affiliated gene expression in a high-NANOG state (Figure 7).

This model of NANOG repression through PRC2 and TET1 does not apply for sporadic genes, such as *Car2*, being specific for lineage-affiliated genes, like *Sox3*. These results are in accordance with preliminary smFISH data from DHenrique lab. Moreover, the results from AA+GSK343 smFISH analysis provided more evidence that PRC2 acts "downstream" of TET1, with GSK343 reverting the AA effect.

Two sortings were performed to generate VNP_L and VNP_H subpopulations. The aim was to fix cells for future smFISH analysis, since it would be possible to focus on cells from low- to high-NANOG state, incubated with GSK343 and/or AA. Thus, the effects on lineage-affiliated gene expression are expected to be more accentuated.

This crosstalk between NANOG, TET1 and PRC2 allows a fast activation of lineage-affiliated gene transcription as soon as cells undergo "lineage priming" in a low-NANOG state. Thus, NANOG is the "master regulator" that imposes the order in the chaos of stochastic gene expression that characterizes pluripotency, by regulating the expression of lineage-affiliated genes.

The beauty of "lineage priming" lies beneath the fact that it is not a deterministic or irreversible process, cells might revert their "decision" by re-expressing NANOG, returning to a more pristine state of pluripotency, or they might "choose" to proceed to an irreversible lineage commitment.

IV. FUTURE WORK

In the future it is essential to analyse by smFISH the fixed FACS-sorted VNP_L and VNP_H cells, exposed to GSK343 and/or AA during 48 hours. If the results are in agreement with the previously described, the model will be consolidated. More cells should be analysed to strengthen the obtained results, since "lineage priming" is a very rare event that only occurs at around 3% of the total population.

Besides, more priming genes (such as *T-Brachyury* for lineage-affiliated gene and *Cldn6* as sporadic gene) should be tested, to know if the model can be generalized to lineage-affiliated genes. Furthermore, a lineage tracing by time-lapse video should be performed to verify the reversibility of "lineage priming".

V. MATERIALS AND METHODS

A. Maintenance of mES Cells

The mES cell lines used in this project were E14tg2a (E14), a gift from Austin Smith's lab, and Nd (for Nanog dynamics). Nd mESCs were derived from E14tg2a and are a BAC-transgenic line for VNP-tagged *Nanog* gene (Abranches *et al.* 2013).

ES cells were routinely grown at 37ºC in a 5% (v/v) CO_2 incubator in "Serum/LIF" conditions: Glasgow Modified Eagles Medium (GMEM 1x, GIBCO) supplemented with 10% (v/v) fetal bovine serum (ES-qualified FBS, Hyclone), 2 ng/mL LIF and 1 mM 2-mercaptoethanol (Sigma), on gelatin-coated 0.1% (v/v) (Sigma) P60 dishes (Nunc). Stem cells were dissociated and re-plated every 48h with fresh medium, at a constant plating density of $3x10^4$ cells/cm^2 with approximately 80-90% of viability (assessed by trypan blue dye exclusion method). Cell morphology was evaluated by an inverted bright field microscope and a mycoplasma test was performed every time cells were frozen.

mES cells were exposed during 48 hours to 1 µM GSK343 (Sigma) and 100 µL/mL AA (Sigma), with replacement for new "Serum/LIF" medium and drugs after 24h.

B. Single-molecule fluorescent in situ hybridization (smFISH)

Nd and E14 mES cells, cultured in "Serum/LIF" conditions, were fixed for smFISH analysis *(Raj et al*. 2006; Guedes *et al*. 2016). After overnight

hybridization with fluorescent Stellaris FISH probes (Biosearch Technologies) for Alexa 594, TMR and Cy5, mES cells were washed twice with PBS and ressuspended in an anti-fade solution.

Cells were imaged within 24 hours on an inverted wide-field fluorescence microscope, the Zeiss Axio Observer, using a 100x 1.4 oil-immersion objective, a ZEISS Axiocam 506 mono camera and filter sets suitable to the excitation and emission wavelengths of the fluorophores used (Alexa 594, Cy5 and TMR). For each image were taken 20 stacks, with 0.3 µm step size each and a total of 80 to 130 positions were acquired. Segmentation of cells and thresholding of mRNA molecules were done using code developed by the Raj Lab (available online) for Matlab software. Finally, statistical analysis was done using the RStudio software.

C. Flow Cytometry (FC) Analysis

For live cells FC analysis and sorting experiments, cells were dissociated every 48 hours and ressuspended in PBS. Nanog:VNP routine analysis was performed in the BD Accuri C6 flow cytometer and the fluorescence activated cell sorting (FACS) experiment was performed in ARIA III cell sorter.

A fraction of cells treated with chemical modulators or DMSO was fixed for intracellular staining - flow cytometry (IC-FC) assay, in PFA, to analyse the percentage of H3K27me3 positive cells upon exposure to GSK343. It was used an anti-H3K27me3 primary antibody (1:200 dilution) and a secondary antibody of anti-α-rabbit Alexa 647 (1:400 dilution).

In each FC acquisition 10.000 events of viable cells were recorded and subsequently analysed using the FlowJo software.

FACS was used to separate the Nd bulk population into low-Nanog:VNP (VNP_L) and high-Nanog:VNP (VNP_H) and its efficiency was assessed immediately after the sorting by analysis of Nanog:VNP expression and of cell viability by the trypan blue dye exclusion method. After the sorting mES cells were replated in “Serum/LIF” conditions supplemented with GSK343 and/or AA, and DMSO in parallel, for 48h, followed by fixation for smFISH.

VI. ACKNOWLEDGMENTS

I would like to thank to my project’s supervisors Prof. Dr. Domingos Henrique and Prof. Dr. Margarida Diogo, and members from DHenrique Lab at Instituto de Medicina Molecular.

VII. REFERENCES

Abranches, E. et al., 2014. Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development (Cambridge, England)*, 141(14), pp.2770–9.

Abranches, E., Bekman, E. & Henrique, D., 2013. Generation and Characterization of a Novel Mouse Embryonic Stem Cell Line with a Dynamic Reporter of Nanog Expression. *PLoS ONE*, 8(3), pp.1–12.

Acampora, D., Di Giovannantonio, L.G. & Simeone, A., 2013. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development (Cambridge, England)*, 140(1), pp.43–55.

Aloia, L., Di Stefano, B. & Di Croce, L., 2013. Polycomb complexes in stem cells and embryonic development. *Development*, 140(12), pp.2525–2534.

Blaschke, K. et al., 2014. Vitamin C induces Tet-dependent DNA demethylation in ESCs to promote a blastocyst-like state. , 500(7461), pp.222–226.

Chambers, I. et al., 2007. Nanog safeguards pluripotency and mediates germline development. *Nature*, 450(7173), pp.1230–4.

Costa, Y. et al., 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, 495(7441), pp.370–4.

Gagliardi, A. et al., 2013. A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *The EMBO journal*, 32(16), pp.2231–47.

Guedes, A.M. V., Henrique, D. & Abranches, E., 2016. Dissecting Transcriptional Heterogeneity in Pluripotency: Single Cell Analysis of Mouse Embryonic Stem Cells. In *Methods in Molecular Biology*. pp. 101–119.

Ito, S. et al., 2010. Role of Tet proteins in 5mC to 5hmC conversion , ES cell self- renewal , and ICM specification. , 466(7310), pp.1129–1133.

Kohli, R.M. & Zhang, Y., 2013. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, 502(7472), pp.472–9.

Martinez Arias, A., Nichols, J. & Schröter, C., 2013. A molecular basis for developmental plasticity in early mammalian embryos. *Development*, 140(17), pp.3499–510.

Neri, F. et al., 2015. TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. *Nucleic acids research*, 43(14), pp.6814–6826.

Osorno, R. & Chambers, I., 2011. Transcription factor heterogeneity and epiblast pluripotency. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366(July), pp.2230–2237.

Raj, A. et al., 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology*, 4(10), pp.1707–1719.

Sui, X. et al., 2012. Crosstalk Between DNA and Histones: Tet’s New Role in Embryonic Stem Cells. *Current Genomics*, 13(8), pp.603–608.

Torres-Padilla, M.-E.E. & Chambers, I., 2014. Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage [review]. *Development (Cambridge, England)*, 141 VN-(11), pp.2173–2181.

Villasante, A. et al., 2011. Epigenetic regulation of Nanog expression by Ezh2 in pluripotent stem cells. *Cell Cycle*, 10(9), pp.1488–1498.

Wu, H. et al., 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*, 473(7347), pp.389–93.

Yin, R. et al., 2013. Ascorbic acid enhances tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *Journal of the American Chemical Society*, 135(28), pp.10396–10403.