UNIVERSITY OF CALIFORNIA SAN DIEGO

Designing Principles for Epigenetic Fluorescence Resonance Energy Transfer Biosensors

A Thesis submitted in partial satisfaction of the

requirements for the degree

Master of Science

in

Bioengineering

by

Ya Gong

Committee in Charge:

Professor Yingxiao Wang, Chair
Professor Prashant Mali
Professor Sheng Zhong

2018

The Thesis of Ya Gong is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2018

Table of Contents

Signature Page iii

Table of Contents iv

List of Figures and Tables v

Acknowledgements vi

Abstract of the Thesis vii

1. Background 1

    1.1 Histone modifications and histone methylation 1

    1.2 Regulation of histone methylation 2

    1.3 Roles of histone methylation in gene transcription 3

    1.4 Histone H3 lysine 27 tri-methylation 4

    1.5 Fluorescence Resonance Energy Transfer Biosensors 4

2. Result 8

    2.1 Construction and testing of the prototype H3K27me3 biosensor 8

    2.2 Construction and testing of YPet-Pc-swapped H3K27me3 biosensors 13

    2.3 Replacing EV linker with shorter-sized linker 14

        2.3.1 Construction and testing of 34mer-linker H3K27me3 biosensor 14

        2.3.2 17mer-linker H3K27me3 biosensor 17

3. Discussion 19

4. Materials & Methods 22

5. References 24

List of Figures and Tables

Table 1: Dissociation constant values (Kd) for the dPC and dHP1 to different methylated histone H3 peptides 6

Figure 1: Functional mechanism of an intramolecular FRET biosensor. 9

Figure 2: Construction and testing of prototype biosensors 11

Figure 3: Construction and testing of YPet-Pc biosensors 14

Figure 4: Construction and testing of 34mer-linker biosensors 16

Figure 5: Construction and testing of 17mer-linker biosensors 18

Figure 6: Schematic representation of the titration curve of FRET/ECFP ratio in intramolecular FRET biosensor 19

Acknowledgements

I would like to acknowledge Dr. Yingxiao Wang for his invaluable guidance and support as the chair of my committee. Also, I would like to acknowledge Dr. Shaoying Lu, Dr. Qin Peng, and Dr. Ziliang Huang, who have worked with me and generously lent their help along this project.

Abstract of the Thesis

Designing Principles for Epigenetic Fluorescence Resonance Energy Transfer Biosensors

by

Ya Gong

Master of Science in Bioengineering

University of California San Diego, 2018

Professor Yingxiao Wang, Chair

Histone proteins in chromatin undergo various modifications that have profound impacts on many cellular processes, including cell cycle control, cancer, senescence, X-inactivation, cell fate decisions, and stem cell differentiation. These histone marks do not occur isolated, but often occur mutually exclusive or concurrently. Despite the extensive research ongoing, a large part of the regulation of histone marks remained elusive due to the lack of powerful and efficient methods. Here, we present a method of constructing an epigenetic fluorescence energy resonance transfer

(FRET) biosensor by tuning variables including the linker length and orientations between the fluorescent proteins in the case of a H3K27me3 FRET biosensor. It reveals that shortening the linker connecting the fluorescent protein pairs and binding partners could indeed increase the FRET change between the bound and unbound states in the biosensor. This key concept can be generally applied to optimize various different FRET biosensors, especially histone epigenetic FRET biosensors with binding domains that have relatively low binding affinities.

1. Background

1.1 Histone modifications and histone methylation

Approximately 2.5 billion base pairs of DNA are packed and stored in a single human diploid cell to provide a physiological template of all eukaryotic genetic information [1, 2]. The huge amount of DNA is compacted into chromatin in the microscopic space of the nucleus with the help of histone proteins. Histone proteins are highly alkaline proteins with positive charges that can associate tightly with the negatively charged DNA and exist in five major families: H1/H5, H2A, H2B, H3, and H4 [3,4]. They are the chief protein in the chromatin around which long sequence of DNA coils, thus playing a pivotal role in regulating the transcription of certain genes. Gene regulation of histone proteins is mainly achieved by the post-translational modification to the histone proteins, including methylation, acetylation, phosphorylation, ubiquitylation, and sumoylation [5]. Post-translational modifications alter the chromatin structure or recruiting histone remodeling enzymes and thus change gene expression. In addition to the regulating gene transcription, post-translation modifications are also known to affect DNA repair, replication and recombination [6].

Among the different kinds of post-translational modifications, methylation on the histone is found to activate or silence the specific genes within the DNA complexed with the methylated histone. Histone methylation mainly occurs on the side chain of lysine and arginine, in which lysine can be mono-. di-, or tri-methylated whereas arginine can only be mono- and di-methylated due to their respective chemical structure [7]. Each location of histone methylation as well as the degree of methylation (mono-, di-, or tri-methylation) have been associated with differential gene expression status. For instance, tri-methylation on histone H3 lysine 4 (H3K4me3) has been associated with active gene transcription [8], whereas tri-methylation on histone H3 lysine 27

(H3K27me3) has been associated with a repressed chromatin state [9]. However, histone methylations do not happen by themselves alone, but are always mutually exclusive and combined methylation marks can lead to different impacts as compared to the same single methylation mark appearing by itself [5]. Although H3K4me and H3K27me3 often come with transcription activation and repression respectively, the presence of both of them simultaneously help poise genes for transcription [10].

1.2 Regulation of histone methylation

Histone methylation is regulated by histone methyltransferases and histone demethylases, by which methylation groups are added to or removed from the histones. Previous research have identified three families of histone methyltransferases that catalyze the addition of methyl group to histones [11]. Proteins that contain SET domain [12] and Dot1 like proteins [13] have been known to methylate lysine, and the PRMT family of proteins have been known to methylate arginine [14]. Two families of demethylases, the amine oxidases and Jumoji C domain containing, iron-dependent dioxygenases, have also been identified to demethylate methyl-lysine, whereas the arginine demethylases remained elusive.

The recruitment of methyltransferase and demethylase onto specific histone targets remained somewhat unclear, but several models have been proposed through previous studies. It has been found that specific DNA sequences can lead to the recruitment of several histone-modifying enzymes. The best studied examples are the Trithorax Group Response Element (TREs) and Polycomb Group Reponse Element (PREs) in *Drosophila*. TREs directs the recruitment of TrX (a H3K4 methyltransferase complex) and PREs directs the recruitment of PcG proteins (the

PRC2 complex catalyzes H3K27me3), possibly through specific DNA binding transcription factors that recognize these elements [15-17].

Long non-coding RNAs (lncRNAs) have also been suggested to target certain methyltransferases and demethylases to specific locations on genome. It has been shown that lnRNAs bind to members of the PRC2 complex [18, 19], the H3K9 methyltransferase G9a [19, 20], and the H3K4 methyltransferase complex member WDR5 [21]. It has been found that the human lncRNA HOTAIR binds to both PRC2 and a demethylase LSD1 containing complex, suggesting that lncRNA plays a role in recruiting an H3K27 methyltransferase and an H3K4 demethylase [22]. Small non-coding RNAs were also found to play a role in directing chromatin modifications, including histone methylation. The RNA interference (RNAi) machinery has been shown to be indispensable for maintaining the centromeric heterochromatin via H3K9 methylation [23]. Moreover, it has been shown in mice that RNAi machinery is linked to the induction of H3K27me3 and heterochromatin changes during the X chromosome inactivation [24].

DNA methylation itself has also been proposed to lead histone methylation. It has been proved that histone methylation further reinforces another DNA methylation to establish a repressive chromatin environment [25, 26]. In *A. thaliana*, the H3K9 methyltransferase SUVH4 was found to bind methylated DNA [27], and mutation of the methyl-DNA binding domain of the Arabidopsis H3K9 methyltransferase SUVH5 decreases H3K9me2 [28].

1.3 Roles of histone methylation in gene transcription

It has been hypothesized that histone methylation could influence transcription through the mechanism of chromosomal looping, bringing physically separate regions of chromatin closer together [29]. However, it remains elusive whether chromosomal looping is a cause or

consequence of transcriptional regulation. Histone modification can also directly or indirectly affect the higher-order chromatin structure via the recruitment of remodeling complexes [30, 31]. Sequence-specific DNA-binding transcription factors, so-called “pioneering factors’, can bind to specific regions on the chromatin and open the inaccessible chromatin domains [32]. Then DNA methylation and histone modifications participate together to make the chromatin more accessible to let in other transcription factors such as RNA polymerase II (RNAPII) and the pre-initiation complex (PIC) [33]. Certain histone methylation patterns are also required for the binding of transcription factors. A study has shown that histone modifications can affect MYC binding to promoters in humans [34]. Though much research has been done to decipher the role of histone methylation in the control of gene expression, more work needs to be done to fully understand the cascade.

1.4 Histone H3 lysine 27 tri-methylation

Histone methylation is known to play important roles in many biological processes, including cell cycle regulation, DNA damage and stress response, development and differentiation [35-38]. Among all forms of histone methylations, we are most interested in the histone H3 lysine 27 tri-methylation (H3K27me3) because it is tightly associated with inactive gene promoters [39], whereas the mono- and di-methylations of H3K27 are less well-studied. Therefore, we will focus on H3K27 tri-methylation in this paper. H3K27me3 only has one known methyltransferase EZH2 [40]. EZH2 is part of the PRC2 complex responsible for repressing many genes involved in development and differentiation [41, 42]. Therefore, it is believed that H3K27me3 is critical for the repression of developmental genes.

PcG proteins were firstly discovered to be essential for the formation and development in *Drosophila melanogaster*. They have been found to impact many cellular processes such as cell cycle control, X-inactivation, stem cell differentiation, cancer and senescence. H3K27me3 is catalyzed by the polycomb group (PcG) protein PRC2 and PRC2 is also known to regulation all forms of H3K27 methylation (mono-, di-, tri-methylation) in a spatially defined manner, contributing to different genomic functions in the cell [43]. The recruiting of PRC2 to specific H3K27 remained elusive, and hypotheses modeling this mechanism have been detailed in section 1.2.

Both PcG proteins PRC1 and PRC2 take part in the gene transcription process through H3K27me3. PRC2 is required to silence the initial target where it binds to, and PRC2 further stabilize this silencing and underlies cellular memory of silenced region after cellular differentiation. Specifically, after being recruited to target H3K27, the EZH2 subunit in PRC2 catalyzes the tri-methylation on H3K27 accompanied by the binding of SUZ12 and EED to the EZH2 subunit [44]. Then the chromodomain in PRC1 binds to the H3K27me3 and propagates the silencing by catalyzing lysine 110 mono-ubiquitylation (K119) of histone H2A, and H2AK119Ub1 contributes to gene silencing through chromatin compaction [45] and inhibition of transcriptional elongation [46].

Chromodomain is highly conserved among human, mice, and fruit flies. In mammalian cells, the eight Cbx1-8 proteins contain the chromodomain. They are evolutionarily related to the *Drosophila* HP1 (*d*HP1) and Pc (*d*Pc) proteins that also contain the conserved chromodomain [47]. *d*HP1 and *d*Pc have been confirmed to specifically recognize repressive marks of H3K9me3 and H3K27me3, respectively (Table 1) [45, 48], whereas the binding specificity of the eight Cbx proteins are more complicated [47]. From the information, we concluded that the *d*PC domain

could be used as a potential binder in our fluorescence resonance energy transfer (FRET) biosensor that can visualize the dynamics H3K27me3 in a single living cell because it's high affinity toward H3K27me3. Visualizing the global as well as locus-specific H3K27me3 dynamics in living cells will render us powerful tools to use for cell reprogramming and developing potential cancer therapies and disease treatments.

Table 1: Dissociation constant values (K_d) for the dPC and dHP1 to different methylated histone H3 peptides. Signal that is too close to the background is designated as n.d. and signa that is too week to give any conclusive result is designated as >1000. Adapted from [49].

K_D (μM)	H3K9ME1	H3K9ME2	H3K9ME3	H3K27ME1	H3K27ME2	H3K27ME3
PC	>1000	>1000	125±28	20±3	28±4	5±1
HP1	46±9	7±2	4±1	n.d.	n.d.	64±7

1.5 Fluorescence Resonance Energy Transfer Biosensors

Fluorescence resonance energy transfer (FRET) is a mechanism describing nonradiative energy transfer between two fluorophores, one donor and one acceptor [50]. This process depends on several factors, such as the spectral overlap of donor fluorophore's emission and acceptor fluorophore's excitation wavelengths, the distance between the fluorophores, and the relative orientation of the fluorophores [50]. Genetically encoded FRET-based biosensors have become powerful tools in visualizing the spatiotemporal dynamics of cellular signaling molecules such as Ca2+, phospholipids, and protein kinases [50, 51] in living cells.

Genetically encoded FRET biosensors are mainly classified into two types: intramolecular FRET biosensor and intermolecular FRET biosensor. The intramolecular FRET biosensor contains

both donor and acceptor fluorophores within a single molecule, whereas the intermolecular FRET biosensor contains a pair of donor and acceptor fluorophore conjugated to two separate molecules respectively. Intramolecular FRET biosensors are more widely used in cell biology due to its advantages of high signal-to-noise ratio, easy loading into the cell, and simple ratiometric image analysis [50, 51]. Figure 1 is an example of a FRET biosensor [52]. On stimulation, the domain binds to the substrate, altering the relative distance or orientation between the connected fluorescent proteins to result in a FRET change [53]. In this way, the molecular signals can be monitored by observing the changes in FRET. Since each cell may have a different copy number of biosensors loaded into them, the change in signal is normalized by dividing FRET with ECFP, and the ratio of FRET/ECFP can be compared prior and after simulation in order to monitor various molecular signals [53].

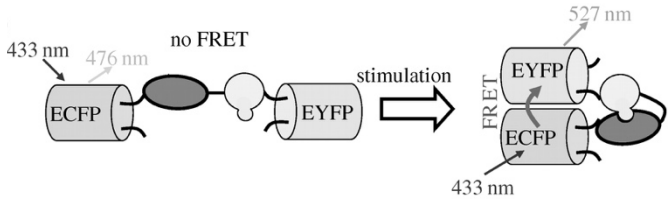

Figure 1: Functional mechanism of an intramolecular FRET biosensor. Adapted from [52]

As previously mentioned, the FRET efficiency of intramolecular biosensors primarily depends on the relative orientation and the distance between the fluorophores [50]. Orientation-dependent biosensors, or biosensors in which the orientation between the two fluorophores plays a dominant role in affecting FRET efficiency, usually exhibit higher sensitivity than the distance-dependent biosensor [53], in which the distance between the fluorophores dominantly changes FRET efficiency. However, since the 3D-structure of the intramolecular biosensor in live cells can hardly

be predicted, time-consuming optimization by trail-and-error are often required in order to optimize the biosensors.

2. Results

2.1 Construction and testing of the prototype H3K27me3 biosensor

To visualize the epigenetic dynamics happening inside living cell, we constructed a fluorescence resonance energy transfer (FRET) biosensor that monitors the tri-methylation at lysine position 27 on histone H3. The prototype construct was assembled based on a previous histone methylation reporter presented by Alice Ting's group dated back to 2004 [54]. However, the reporter published by Ting's group lacked in vivo testing, had a low FRET contrast between methylated and non-methylated states, and could not efficiently localize into the nucleus after our testing. Based on this paper, we built a new prototype construct with the same truncated *Drosophila* Polycomb (Pc) chromodomain as the binding domain and a full-length histone H3 containing the lysine K27 position (Figure 2A). The Pc was truncated to include residue 21 to 78 of the full Pc domain, which includes the aromatic rings necessary for the specific binding to H3K27me3 [55]. The affinity of the Pc domain to H3K27 tri-methylation is known to be 5 μM [49]. To further improve nuclear localization and simulation of cellular events, a full-length histone H3 containing lysine 27 was used as the substrate-containing domain. The fluorescent pairs were changed to ECFP and YPet, which are fluorescent donor and acceptor optimized for FRET, respectively. The ECFP and YPet were sandwiched between the binding partners so as to increase the chance of FRET when binding event happens. A 120-amino-acid-long EV linker with repetitions of amino acids – glycine, serine and alanine to evade extra background FRET was used to link ECFP and YPet. Since the 3D conformation of the biosensors in cells were unknown, we

made two different prototype constructs. The GGS-linker prototype biosensor has a 15-mer GGS-linker separating Pc and YPet to provide enough flexibility which may lead to better FRET efficiency upon binding. The no-linker prototype biosensor does not have linker between Pc and YPet. It is unclear whether the linker, upon binding would provide flexibility for better FRET or limits FRET with too much flexibility. Therefore, we tested both inside the HEK cells.

We hypothesize that when there is no tri-methylation on H3K27, the biosensor sits in the "OFF" state, with the fluorescent proteins relatively far away from each other and no FRET occurs. When H3K27 is tri-methylated, Pc binds to H3K27me3 and induces a global conformational change in the biosensor, brings the YPet close to ECFP and causing FRET. Biosensor is stimulated to an "ON" state. When the methylated H3K27 is further demethylated, Pc releases non-methylated H3K27 and returns to "OFF" state with no FRET again (Figure 2B). By measuring the ratio of FRET/ECFP, we are able to determine the global tri-methylation level on H3K27.

After successfully constructing the biosensors, we tested them inside HEK cells via lipofectamine 3000 transfection and looked at FRET after 48 hours. In order to validate that FRET only changes when H3K27 is tri-methylated, two different mutants at H3K27 were made. The L mutation mutated H3 lysine 27 to leucine, whose side chain cannot be methylated. The M mutation mutated H3 lysine 27 to methionine, which previously had been shown to inhibit global level of H3K27 tri-methylation in cells by inhibiting the SET-domain methyltransferases [55].

It was observed that all the constructs were able to locate specifically into the cell nucleus and the constructs did not have any significant effect on the viability and conditions of the cells (Figure 2C, 2D, 2H, 2I). The FRET/ECFP ratios in HEK cells transfected with the same biosensor were averaged and the averages of different groups were statistically analyzed using One-way Anova. The results showed that in the GGS-linker biosensors, the wildtype biosensor has a

significantly higher FRET ratio than the L mutated biosensor, and the L mutated one has a significantly higher FRET ratio than the M mutated biosensor (Figure 2C). A same trend was also found in the no-linker biosensors (Figure 2D). This was somewhat as we expected, because L mutants could only inhibit the binding of Pc domain onto the H3K27 on the biosensor, whereas M mutants inhibits the tri-methylation on both endogenous H3K27 and the H3K27 on biosensor. These results confirmed that our wildtype biosensor could specifically detect H3K27 tri-methylation.

From the results, we also noticed that the no-linker biosensors overall had lower basal FRET ratios compared to the GGS-linker ones, suggesting the GGS-linker between Pc and YPet may have provided the flexibility necessary for better FRET in presence of H3K27me3. With the same set of data, we further compared the averages between the GGS-linker biosensors and no-linker biosensors within the wildtype and L, M mutations. It was found that the GGS-linker wildtype biosensor had significantly higher FRET ratio than the no-linker wildtype biosensors (Figure 2E). In the L-mutated biosensors, the GGS-linker biosensor also had significantly higher FRET ratio than the no-linker biosensor, but the significance became smaller (Figure 2F). In the M-mutated biosensors, there is no significance in FRET ratios between the GGS-linker and no-linker biosensors (Figure 2G). This is consistent with our previous results, because in L mutation, there is much less tri-methylation on H3K27, and both biosensors had very low chance of binding; in M mutation, most of tri-methylation on H3K27 is inhibited, and both biosensors could not bind at all. By comparing the GGS-linker biosensors with no-linker biosensors, we discovered that the GGS-linker biosensors were more sensitive than the no-linker biosensors.

To further find out that the biosensors work in a way as we expected, we mutated the Pc binding domain. A mutation on Y26 to lysine was introduced to the Pc domain to abolish one of

the aromatic ring side chains that is crucial to the specific recognition and binding to H3K27me3. Previous literature also confirmed that Y26K mutation significantly reduce the binding between

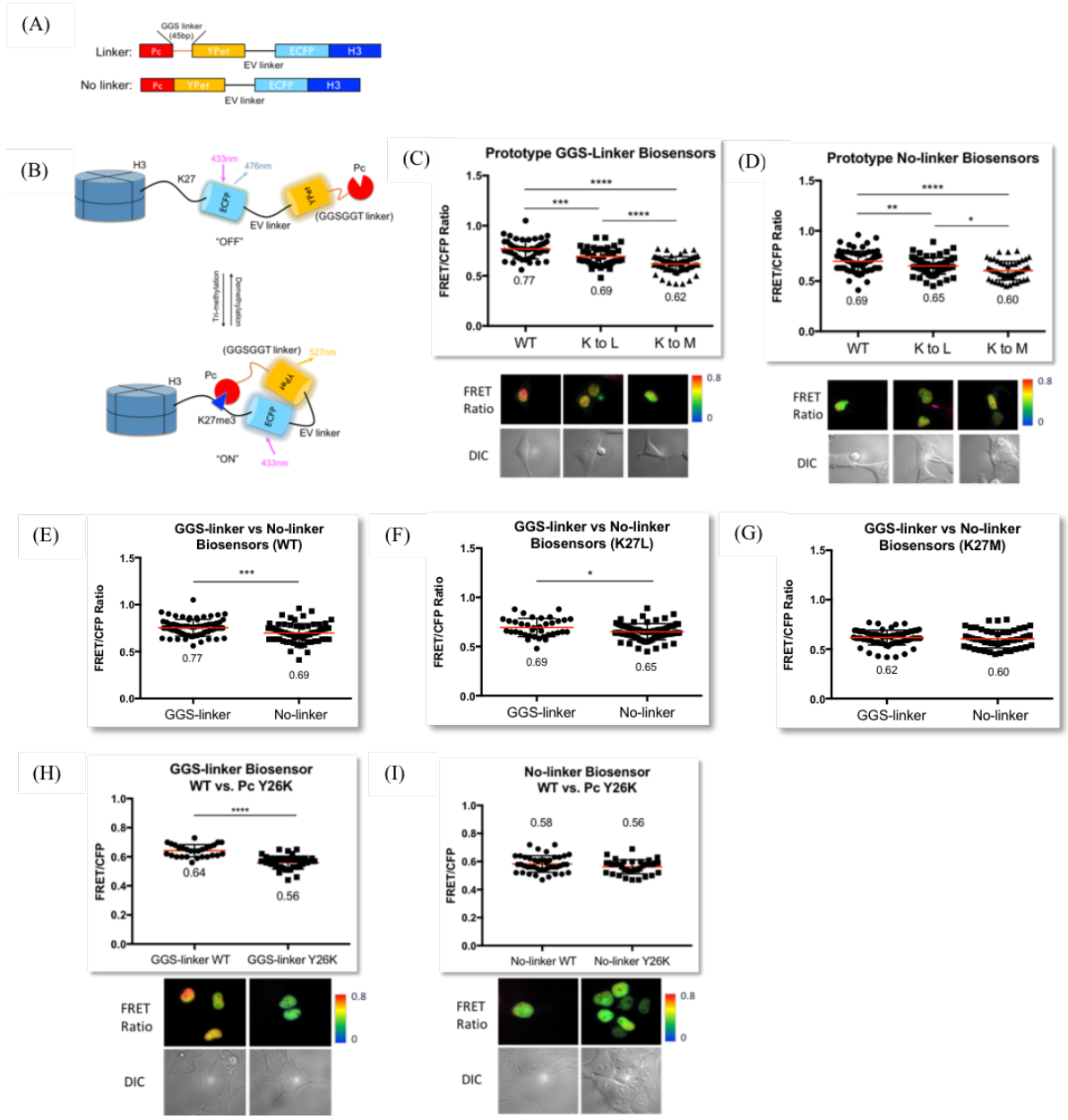

Figure 2: Construction and testing of prototype biosensors. (A) Maps of the GGS-linker prototype H3K27me3 biosensor and no-linker prototype H3K27me3 biosensor. (B) Schematic presentation of the FRET mechanism of the GGS-linker prototype biosensor in cell. (C) The basal FRET/ECFP ratios and images of the wildtype and H3K27L, H3K27M mutations in GGS-linker prototype H3K27me3 biosensors. (D) The basal FRET/ECFP ratios and images of the wildtype and H3K27L, H3K27M mutations in no-linker prototype H3K27me3 biosensors. (E) Comparison of the FRET/ECFP ratios between the GGS-linker and no-linker prototype H3K27me3 biosensors in wildtype. (F) Comparison of the FRET/ECFP ratios between the GGS-linker and no-linker prototype H3K27me3 biosensors in K27L mutation. (G) Comparison of the FRET/ECFP ratios between the GGS-linker and no-linker prototype H3K27me3 biosensors in K27M. (H) The basal FRET/ECFP ratios and images of the wildtype and Pc Y26K mutated GGS-linker prototype H3K27me3 biosensors. (I) The basal FRET/ECFP ratios and images of the wildtype and Pc Y26K mutated no-linker prototype H3K27me3 biosensors.

Pc and H3K27me3 [49]. In the GGS-linker biosensors, we saw a significantly higher FRET ratio

in the wildtype GGS-linker biosensor compared to the Y26K mutated one, confirming that Pc binds to H3K27me3 as we expected, and the previous results of the GGS-linker biosensors are reliable (Figure 2H). However, in the no-linker biosensors, there was no difference between the wildtype and Y26K mutated one, suggesting that the Pc in no-linker biosensors may bind to H3K27me3 in an unexpected mechanism, or that the previous results from the no-linker biosensors are false positive (Figure 2I). Based on all the results from testing the GGS-linker and no-linker prototype biosensors, we are confident that the GGS-linker biosensor is working as we expected and can specifically reflect H3K27me3 level in HEK cells. However, the contrast should be improved to provide a more reliable and sensitive tool.

The efficiency of FRET mainly depends on the relative orientation and distance between the fluorophores. In our case, we first tried to make the biosensor a "distance-dependent" one, in which the distance between ECFP and YPet is the main factor that changes FRET. However, the FRET gain was low with and without stimulation. Therefore, we will focus on changing the orientation between ECFP and YPet to enhance FRET efficiency when H3K27me3 is present. We can improve the FRET efficiency through changing orientation in two different ways. One is to switch the positions of Pc domain and YPet for a different orientation, and the other is to shorten the EV linker to bring the two fluorescent proteins closer to each other, changing the biosensor from a "distance-dependent" type to an "orientation-dependent" type.

2.2 Construction and testing of YPet-Pc-swapped H3K27me3 FRET Biosensors

The results from testing the prototype H3K27me3 FRET biosensors gave us a relatively small contrast between the wildtype and the mutant biosensors, indicating an unstable binding between the Pc domain and H3K27me3. It is impossible to know the real conformational change of the biosensor inside living cell upon Pc domain and H3K27me3 binding, thus we tried swapping the position of YPet and Pc domains to place YPet at the C-terminal and Pc inside in hope to achieve better interaction between the fluorescent pair on binding between Pc domain and H3K27me3 (Figure 3A). However, after testing the constructs in HEK cells under the same conditions, we found that the significant difference between the wildtype and H3K27 mutants we saw in our prototype were completely abolished. The FRET ratio of the wildtype linker construct also decreased from a range between 0.6-0.7 to a range between 0.50-0.53. Since the FRET ratios of all the biosensors, whether mutated or not, had very similar FRET ratio in the range of 0.50-0.53, it is unlikely that the signals were false signals (Figure 3B, 3C). Therefore, we concluded that the YPet-Pc swapped biosensor could not be used to reflect the tri-methylation on H3K27 in cells. Bringing the YPet to the C terminal might have caused YPet and ECFP to be away from each other, no matter the binding events happens or not. The results also confirmed that our original prototype design with the fluorescent pairs sandwiched between the Pc binding domain and H3 should be the best positions to allow most interaction between the fluorescent pairs upon binding event.

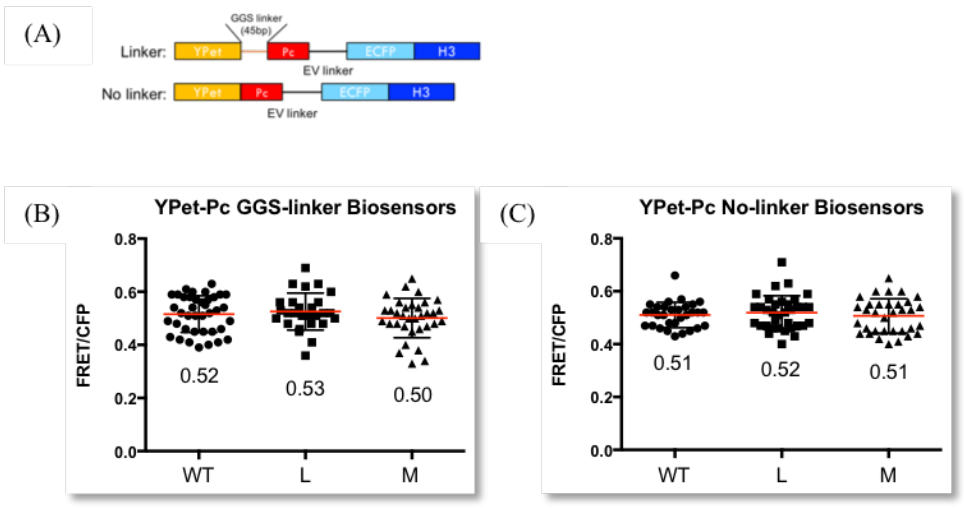

Figure 3: Construction and testing of YPet-Pc biosensors. (A) Maps of the GGS-linker YPet-Pc H3K27me3 biosensor and no-linker YPet-Pc H3K27me3 biosensor. (B) The basal FRET/ECFP ratios of the wildtype and H3K27L, H3K27M mutated GGS-linker biosensors. (C) The basal FRET/ECFP ratios of the wildtype and H3K27L, H3K27M mutated no-linker biosensors.

2.3 Replacing EV linker with shorter linker

Since switching the positions of Pc domain and YPet did not work as expected, therefore, we moved onto shortening the length of the EV linker. A shorter linker between the fluorescent proteins will make the biosensor into an "orientation-dependent" type and may lead to a more favorable conformation for FRET. Shortening the linker can also bring the binding partners closer to each other and thus increase the chance for binding by increasing their local concentrations. However, a shorter linker also indicates a higher possibility of FRET even when no binding event occurs, or higher background signal. Thus, to optimize the FRET ratio change, the linker length should be optimized to allow higher contribution of the binding and lower contribution of the background to the FRET change. In this purpose, we shortened the 120mer EV linker into a 34mer-linker and a 17mer-linker with everything else remain the same as the prototype biosensors.

2.3.1. Construction and testing of 34mer-linker H3K27me3 biosensors

In the 34mer biosensors, a linker that is 34-amino-acid-long linker with sequence "GSTSGSGKPGSGEGSTKGSTSGSGKPGSGEGSTK" was used to replace the EV liker in both GGS-linker prototype biosensors and no-linker prototype biosensors (Figure 4A). After testing the 34mer-linker biosensors in HEK via lipofectamine 3000 transfection, it was found that all the basal ratios elevated from an average of 0.6 to an average of 1.3 (Figure 4B, 4C). This is consistent with our expectation that shortening the EV linker can increase the interaction between the fluorescent pairs and thus FRET overall all, regardless whether binding happens or not. However, the contrast between wildtype and mutant biosensors were a little complicated. In the GGS-linker biosensors, it was found that the wildtype biosensor had significantly higher FRET ratio than the H3K27L, H3K27M and Pc domain Y26K mutants (Figure 4B). In the no-linker group, there is no significant

difference between the wildtype and H3K27L and H3K27M mutant biosensors at all, but the wildtype did have significantly higher FRET ratio than the Pc Y26K mutated biosensor (Figure 4C). This implies that the no-linker biosensor with a 34mer linker worked in the mechanism that Pc binds to H3 as we expected, but it may not be sensitive to differentiate among wildtype H3K27me3, H3K27L, and H3K27M.

Surprisingly, after shortening the linker to nearly one-fourth of the original EV linker, the largest FRET contrast, which is found between wildtype GGS-linker biosensor and Pc Y26K mutated GGS-linker biosensor, decreased to 14%. It seems that the 34mer linker indeed increased the background FRET, but was not able to enhance the FRET gain after tri-methylation at H3K27.

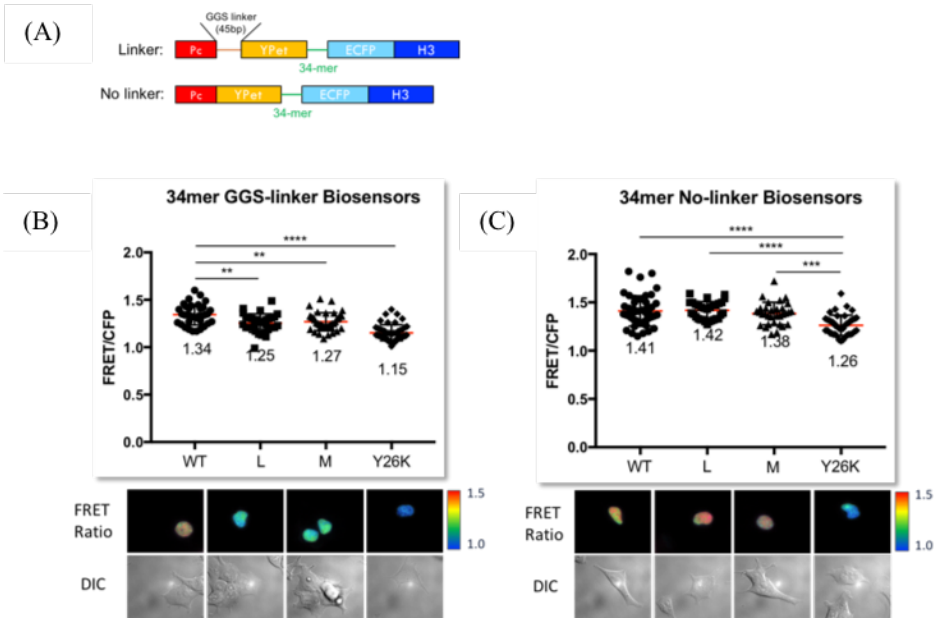

Figure 4: Construction and testing of 34mer-linker biosensors. (A) Maps of the GGS-linker 34mer-linker H3K27me3 biosensor and no-linker 34mer-linker H3K27me3 biosensor. (B) The basal FRET/ECFP ratios of the wildtype, H3K27L, H3K27M, and Pc Y26K mutated GGS-linker biosensors. (C) The basal FRET/ECFP ratios of the wildtype and H3K27L, H3K27M, and Pc Y26K mutated no-linker biosensors.

2.3.2. Construction and testing of 17mer-linker H3K27me3 biosensors

The EV linker was further shortened to a 17-amino-acid-long linker with sequence "GSTSGSGKPGSGEGSTK" (Figure 5A). After testing the 17mer-linker biosensors in HEK via lipofectamine 3000 transfection, we saw that the basal FRET ratios of all the biosensors again increased to the range of 1.7-5.0, which was as we expected and consistent with previous results of the 34mer-linker biosensors (Figure 5B, 5C). Moreover, the results showed that the wildtype GGS-linker biosensor had a significantly higher FRET ratio than all the mutants, including H3K27L, H3K27M, and Pc Y26K (Figure 5B). A similar trend can also be found in the no-linker biosensors (Figure 5C). No significant difference was found between the L and M mutants, which was also as we expected because in neither case can the H3K27 on the biosensor be methylated. It also suggests that the Pc domain on the biosensor only binds to the H3 on the biosensor instead of other endogenous H3 tails due to the shortening of the linker and thus confined conformation.

It is also worth noticing that the wildtype GGS-linker and no-linker biosensors both have a wider distribution of FRET ratios across the sample cells compared to the prototype biosensors and 34mer-linker biosensors. This may be a hint of an optimized biosensor because the 17mer-linker wildtype biosensors showed a wider dynamic range. The wider distribution may also suggest that the 17mer-linker has a higher sensitivity than the previous biosensors because H3K27me3 level in cells is dependent on different stages of the cell cycle [], and this wider distribution may have reflected this phenomenon, since all the cells were randomly selected to be imaged. Overall, the largest contrast, which was again found between the GGS-linker wildtype biosensor and GGS-linker Pc Y26K mutated biosensor, increased to 32%.

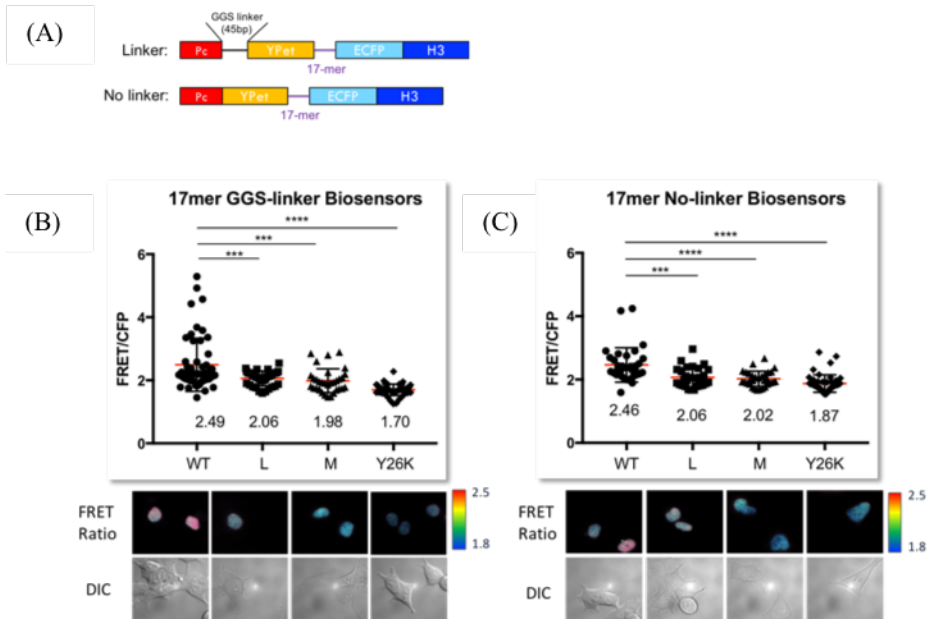

Figure 5: Construction and testing of 17mer-linker biosensors. (A) Maps of the GGS-linker 17mer-linker H3K27me3 biosensor and no-linker 17mer-linker H3K27me3 biosensor. (B) The basal FRET/ECFP ratios of the wildtype, H3K27L, H3K27M, and Pc Y26K mutated GGS-linker biosensors. (C) The basal FRET/ECFP ratios of the wildtype and H3K27L, H3K27M, and Pc Y26K mutated no-linker biosensors.

3. Discussion

In FRET biosensors, the dynamic range of the biosensor is defined as the range of the FRET/ECFP ratio between ON and OFF states of the biosensor (Figure 6). The gain of the FRET signal is defined as the relative change in FRET/CFP ratio before and after stimulation. In practice, the activity and the concentration of the target molecule can be monitored by observing the changes in FRET/ECFP before and after stimulation. Therefore, the gain of a FRET biosensor depends both on the dynamic range of the FRET biosensor as well as the relative increase after stimulation (ON state) versus that before stimulation (Figure 6). Meanwhile, the sensitivity of FRET biosensors is defined as a concentration of stimulants that increases the FRET/ECFP ratio to 50% of the dynamic range (Figure 6) [56]. In our 17mer-linker biosensors, we were able to see a wider distribution in the wildtype biosensors, indicating that they both have a larger dynamic range compared to the previous prototype biosensors, and thus, a higher sensitivity.

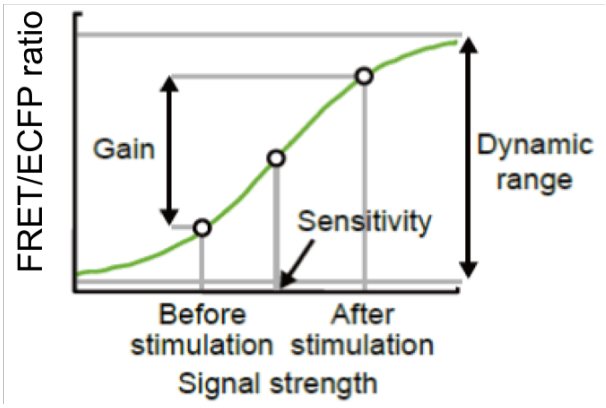

Figure 6: Schematic representation of the titration curve of FRET/ECFP ratio in intramolecular FRET biosensor. Adapted from [56].

In addition to the dynamic range, the gain of a FRET biosensor also depends on the relative FRET ratio change between ON and OFF states of the biosensors and this relationship can be expressed as:

$$FRET\ gain = \frac{\text{“ON”}}{\text{“OFF” } (background)}.$$

In our case, by shortening the EV linker, we increased both the FRET ratio of the biosensor in the OFF state as well as the FRET ratio in the ON state. In the case of 34mer-linker biosensors, the increase in OFF state, or background signal is larger than that in the ON state, therefore, resulting in a smaller FRET gain than previously. However, in the case of 17mer-linker biosensors, the increase in ON state FRET ratio is larger than the increase in the background FRET ratio. Thus, we were able to get a higher FRET gain, and eventually, a higher sensitivity.

A previous paper used a longer EV linker and successfully optimized several kinase FRET biosensors, such as the ERK and PKA biosensors [56]. In their case, the binding affinities between the substrate molecule and binding domains on the kinase FRET biosensors are usually very strong [57, 58]. Therefore, their strategy to optimize such kinase FRET biosensors was to lengthen the linker separating the two binding domains and fluorescent proteins, making the biosensors completely “distance-dependent”, so that the background FRET can be lowered to increase FRET gain. Since binding affinities are high, background FRET signal is the dominant factor that affects FRET gain. However, in our case, due to the relatively low binding affinity between Pc domain and H3K27me3, the ON FRET signal becomes the dominant factor that determines FRET gain. Changing it to an “orientation-dependent” type of biosensor by shortening the linker can help increase the FRET gain, because when the two fluorophores are brought closer, the Pc domain and H3K27me3 are also closer, resulting a larger possibility that they can reach each other and bind. That is, the local concentrations of the binding partners increased. Indeed, by shortening the linker to about one seventh of the original length, we were able to see a higher FRET gain. Though the background FRET signal was also raised, as long as the FRET signal increase in the ON state is larger than that in the OFF state, the biosensor’s FRET gain can be more optimized.

There is another thing that worth noticing when we compare the 34mer-linker biosensors and the 17mer-linker biosensors. We found that shortening to a 34mer-linker was not able to increase the FRET gain, whereas shortening the linker further to a 17mer-linker was able to increase the FRET gain. This is very interesting because the relationship between the linker length and FRET gain is not linear/proportional. Shortening to a 34mer-linker may have increased the background FRET signal to a larger extent than increasing the FRET in ON state. Therefore, some trial-and-error is still needed in adjusting the linker size. But compared to the usual trial-and-error optimizations without much directions, this key concept can be applied to improve the sensitivities of many other FRET biosensors, especially the epigenetic FRET biosensors, since they all tend to have binding domains with relatively low affinities.

In conclusion, in our design of this epigenetic biosensors, we found out that playing with the linker length as well as the binding affinity we are able to enhance the FRET contrast between the non-binding and binding of Pc domain and H3K27me3.

4. Materials & Methods

4.1. DNA construction and plasmids

All the constructs were first amplified to obtain necessary fragments through standard polymerase chain reaction (PCR). Then the components were assembled into pcDNA3.1+ vector through Gibson Assembly [New England Biolabs]. All plasmids were purified with the QIAquick Gel Extraction Kit [Qiagen] and Miniprep Kit [Qiagen].

4.2 Cell culture and reagents

HEK293 cells were obtained from American Tissue Culture Collection (ATCC) (Manassas, VA), and they were cultured in medium containing Dulbecco's modified Eagle medium (DMEM) (Gibco), 10% fetal bovine serum (FBS) (Atlanta Biologicals, Lawrenceville, GA), and 1x penicillin/streptomycin (Invitrogen). They were cultured in 37°C humidified-incubator with 5% CO_2. HeLa cells were transfected with the plasmids using Lipofectamine 3000 kit (Sigma-Aldrich) 48 hours before imaging.

4.3 Image acquisition and analysis

Before imaging, the cells were plated onto glass-bottom dishes (Cell E&G) coated with fibronectin (Sigma) at 10 μg/mL concentration, and the coated dishes were kept at 37°C for 4 hours before cell plating. During imaging, the plated HEK cells were maintained in the starvation medium at 37°C with CO2 supplemented.

The images were collected with a Nikon eclipse Ti inverted microscope with 100x DIC Nikon microscope objective (NA 1.4) used and the MetaMorph 7.8.8.0 software (Molecular Devices). Furthermore, the microscope was installed with a 450DRLP dichroic mirror, a

420DF20 excitation filter, 475DF40 filter for ECFP and 535DF25 filter for YPet (Chroma). Image analysis was conducted on MetaMorph 7.8.8.0 software, where background signal was subtracted, and the FRET ratio was recorded as an average value over the whole nucleus.

5. References

1. Jenuwein, T. Translating the Histone Code. *Science* 293, 1074-1080 (2001).
2. Annunziato, A. DNA Packaging: Nucleosomes and Chromatin. *Nature Education* 1(1):26 (2008)
3. Van Holde, K. Chromatin. (Springer-Verlag, 1988).
4. Lehninger, A., Nelson, D. & Cox, M. Lehninger principles of biochemistry. (W.H. Freeman, 2005).
5. Bannister, A. & Kouzarides, T. Regulation of chromatin by histone modifications. Cell Research 21, 381-395 (2011).
6. Xhemalce B, Dawson MA, Bannister AJ. Histone modifications in: Meyers R, ed. Encyclopedia of Molecular Cell Biology and Molecular Medicine. (John Wiley and Sons, 2011).
7. Ng, S., Yue, W., Oppermann, U. & Klose, R. Dynamic protein methylation in chromatin biology. Cellular and Molecular Life Sciences 66, 407-422 (2008).
8. Santos-Rosa, H., Schneider, R., Bannister, A., Sherriff, J., Bernstein, B., Emre, N., Schreiber, S., Mellor, J. & Kouzarides, T. Active genes are tri-methylated at K4 of histone H3. Nature419, 407-411 (2002).
9. Heintzman, N., Stuart, R., Hon, G., Fu, Y., Ching, C., Hawkins, R., Barrera, L., Van Calcar, S., Qu, C., Ching, K., Wang, W., Weng, Z., Green, R., Crawford, G. & Ren, B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature Genetics 39, 311-318 (2007).
10. Bernstein, B., Mikkelsen, T., Xie, X., Kamal, M., Huebert, D., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. & Lander, E. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell 125, 315-326 (2006).
11. Murray, K. The Occurrence of iε-N-Methyl Lysine in Histones. Biochemistry 3, 10-15 (1964).
12. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B., Sun, Z., Schmid, M., Opravil, S., Mechtler, K., Ponting, C., Allis, C. & Jenuwein, T. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-599 (2000).
13. Bannister, A. & Kouzarides, T. Regulation of chromatin by histone modifications. Cell Research 21, 381-395 (2011).
14. Feng, Q., Wang, H., Ng, H., Erdjument-Bromage, H., Tempst, P., Struhl, K. & Zhang, Y. Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain. Current Biology 12, 1052-1058 (2002).
15. Chan, C. S., Rastelli, L. & Pirrotta, V. A Polycomb response element in the *Ubx* gene that determines an epigenetically inherited state of repression. EMBO J. 13, 2553–2564 (1994).
16. Tillib, S. et al. Trithorax- and Polycomb-group response elements within an Ultrabithorax transcription maintenance unit consist of closely situated but separable sequences. Mol. Cell. Biol. 19, 5189–5202 (1999).
17. Fritsch, C., Brown, J. L., Kassis, J. A. & Muller, J. The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a *Drosophila* homeotic gene. Development 126, 3905–3913 (1999).

18. Rinn, J., Kertesz, M., Wang, J., Squazzo, S., Xu, X., Brugmann, S., Goodnough, L., Helms, J., Farnham, P., Segal, E. & Chang, H. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. Cell 129, 1311-1323 (2007).

19. Pandey, R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-DiNardo, D. & Kanduri, C. Kcnq1ot1 Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation. Molecular Cell 32, 232-246 (2008).

20. Nagano, T., Mitchell, J., Sanz, L., Pauler, F., Ferguson-Smith, A., Feil, R. & Fraser, P. The Air Noncoding RNA Epigenetically Silences Transcription by Targeting G9a to Chromatin. Science 322, 1717-1720 (2008).

21. Wang, K., Yang, Y., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B., Protacio, A., Flynn, R., Gupta, R., Wysocka, J., Lei, M., Dekker, J., Helms, J. & Chang, H. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472, 120-124 (2011).

22. Tsai, M., Manor, O., Wan, Y., Mosammaparast, N., Wang, J., Lan, F., Shi, Y., Segal, E. & Chang, H. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. Science 329, 689-693 (2010).

23. Verdel, A. RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. Science 303, 672-676 (2004).

24. Ogawa, Y., Sun, B. & Lee, J. Intersection of the RNA Interference and X-Inactivation Pathways. Science 320, 1336-1341 (2008).

25. Fuks, F. DNA methylation and histone modifications: teaming up to silence genes. Current Opinion in Genetics & Development 15, 490-495 (2005).

26. Rountree, M. & Selker, E. DNA methylation and the formation of heterochromatin in Neurospora crassa. Heredity 105, 38-44 (2010).

27. Johnson, L., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J. & Jacobsen, S. The SRA Methyl-Cytosine-Binding Domain Links DNA and Histone Methylation. Current Biology 17, 379-384 (2007).

28. Rajakumara, E., Law, J., Simanshu, D., Voigt, P., Johnson, L., Reinberg, D., Patel, D. & Jacobsen, S. A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. Genes & Development 25, 137-152 (2011).

29. Dekker, J. Capturing Chromosome Conformation. Science 295, 1306-1311 (2002).

30. Suganuma, T. & Workman, J. Signals and Combinatorial Functions of Histone Modifications. Annual Review of Biochemistry 80, 473-499 (2011).

31. Bell, O., Tiwari, V., Thomä, N. & Schübeler, D. Determinants and dynamics of genome accessibility. Nature Reviews Genetics 12, 554-564 (2011).

32. Cirillo, L., Lin, F., Cuesta, I., Friedman, D., Jarnik, M. & Zaret, K. Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. Molecular Cell 9, 279-289 (2002).

33. Serandour, A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Metivier, R., Salbert, G. & Eeckhoute, J. Epigenetic switch

involved in activation of pioneer factor FOXA1-dependent enhancers. Genome Research 21, 555-565 (2011).

34. Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L. & Amati, B. Myc-binding-site recognition in the human genome is determined by chromatin context. Nature Cell Biology 8, 764-770 (2006).
35. Pedersen, M. & Helin, K. Histone demethylases in development and disease. Trends in Cell Biology 20, 662-671 (2010).
36. Nottke, A., Colaiacovo, M. & Shi, Y. Developmental roles of the histone lysine demethylases. Development 136, 879-889 (2009).
37. Greenberg, R. Histone tails: Directing the chromatin response to DNA damage. FEBS Letters 585, 2883-2890 (2011).
38. Eissenberg, J. & Shilatifard, A. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Developmental Biology 339, 240-249 (2010).
39. Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I. & Zhao, K. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129, 823-837 (2007).
40. Kuzmichev, A. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes & Development 16, 2893-2905 (2002).
41. Boyer, L., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L., Lee, T., Levine, S., Wernig, M., Tajonar, A., Ray, M., Bell, G., Otte, A., Vidal, M., Gifford, D., Young, R. & Jaenisch, R. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441, 349-353 (2006).
42. Bracken, A. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes & Development 20, 1123-1136 (2006).
43. Ferrari, K., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T. & Pasini, D. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. Molecular Cell 53, 49-62 (2014).
44. Morey, L. & Helin, K. Polycomb group protein-mediated repression of transcription. Trends in Biochemical Sciences 35, 323-332 (2010).
45. Pasini, D., Bracken, A., Jensen, M., Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. The EMBO Journal 23, 4061-4071 (2004).
46. Francis, N. Chromatin Compaction by a Polycomb Group Protein Complex. Science 306, 1574-1577 (2004).
47. Ferrari, K., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T. & Pasini, D. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. Molecular Cell 53, 49-62 (2014).
48. Morey, L. & Helin, K. Polycomb group protein-mediated repression of transcription. Trends in Biochemical Sciences 35, 323-332 (2010).
49. Fischle, W. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes & Development 17, 1870-1881 (2003).

50. Miyawaki, A. Visualization of the Spatial and Temporal Dynamics of Intracellular Signaling. Developmental Cell 4, 295-305 (2003).

51. Aoki, K., Kiyokawa, E., Nakamura, T. & Matsuda, M. Visualization of growth signal transduction cascades in living cells with genetically encoded probes based on Forster resonance energy transfer. Philosophical Transactions of the Royal Society B: Biological Sciences 363, 2143-2151 (2008).

52. Liu, B., Kim, T. & Wang, Y. Live cell imaging of mechanotransduction. Journal of The Royal Society Interface 7, S365-S375 (2010).

53. Jares-Erijman, E. & Jovin, T. FRET imaging. Nature Biotechnology 21, 1387-1395 (2003).

54. Lin, C., Jao, C. & Ting, A. Genetically Encoded Fluorescent Reporters of Histone Methylation in Living Cells. Journal of the American Chemical Society 126, 5982-5983 (2004).

55. Lewis, P., Muller, M., Koletsky, M., Cordero, F., Lin, S., Banaszynski, L., Garcia, B., Muir, T., Becher, O. & Allis, C. Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. Science 340, 857-861 (2013).

56. Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y. & Matsuda, M. Development of an optimized backbone of FRET biosensors for kinases and GTPases. Molecular Biology of the Cell 22, 4647-4656 (2011).

57. Felder, S., Zhou, M., Hu, P., Ureña, J., Ullrich, A., Chaudhuri, M., White, M., Shoelson, S. & Schlessinger, J. SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. Molecular and Cellular Biology 13, 1449-1455 (1993).

58. Durocher, D., Taylor, I., Sarbassova, D., Haire, L., Westcott, S., Jackson, S., Smerdon, S. & Yaffe, M. The Molecular Basis of FHA Domain: Phosphopeptide Binding Specificity and Implications for Phospho-Dependent Signaling Mechanisms. Molecular Cell 6, 1169-1182 (2000).