




RESEARCH NOTE

REVISED DNA sequence features in the establishing of H3K27ac

[version 2; referees: 2 approved]

Anatoliy Zubritskiy ¹, Yulia A. Medvedeva¹⁻³

¹Institute of Bioengineering, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, 119071, Russian Federation

²Department of Computational Biology, Vavilov institute of General Genetics, Russian Academy of Sciences, Moscow, 119991, Russian Federation

³Department of Biological and Medical Physics, Moscow Institute of Physics and Technology, Dolgoprudny, 141701, Russian Federation

v2 First published: 08 Feb 2018, 7:165 (doi: [10.12688/f1000research.13441.1](https://doi.org/10.12688/f1000research.13441.1))
 Latest published: 09 Aug 2018, 7:165 (doi: [10.12688/f1000research.13441.2](https://doi.org/10.12688/f1000research.13441.2))

Abstract

The presence of H3K27me3 has been demonstrated to correlate with the CpG content. In this work, we tested whether H3K27ac has similar sequence preferences. We performed a translocation of DNA sequences with various properties into a beta-globin locus to control for the local chromatin environment. Our results suggest that in contrast to H3K27me3, H3K27ac gain is unlikely affected by the CpG content of the underlying DNA sequence, while extremely high GC-content might contribute to the gain of the H3K27ac.

Keywords

histone modification, H3K27ac, GC-content, CpG dinucleotides

Open Peer Review

Referee Status:  

Invited Referees

1 2

REVISED

version 2

published
09 Aug 2018

version 1

published
08 Feb 2018



report



report

1 **Vasily N Aushev** , Icahn School of Medicine at Mount Sinai, USA

2 **Oleg Gusev** , RIKEN, Japan

Discuss this article

Comments (0)

Corresponding author: Yulia A. Medvedeva (ju.medvedeva@gmail.com)

Author roles: **Zubritskiy A:** Investigation, Methodology, Validation, Visualization, Writing – Original Draft Preparation; **Medvedeva YA:** Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Russian Science Foundation [15-14-30002].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Zubritskiy A and Medvedeva YA. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Zubritskiy A and Medvedeva YA. **DNA sequence features in the establishing of H3K27ac [version 2; referees: 2 approved]** *F1000Research* 2018, **7**:165 (doi: [10.12688/f1000research.13441.2](https://doi.org/10.12688/f1000research.13441.2))

First published: 08 Feb 2018, **7**:165 (doi: [10.12688/f1000research.13441.1](https://doi.org/10.12688/f1000research.13441.1))

REVISED Amendments from Version 1

We carefully checked all the names of the genes and plasmids as well as the rest of the text for typos and modified a few statements.

See referee reports

Introduction

Modification of histone proteins is a key mechanism of epigenetic regulation. Histone modifications vary between cells in some genomic locations but not others. This observation raises the question to what extent histone modifications depend on the underlying nucleotide sequences. It has been reported that the attraction of PRC2 complex and consequent H3K27me3 is positively correlated with the local density of CpG dinucleotides¹. More complex sequence patterns, such as transcription factor (TF) binding sites (TFBS), also affect the presence of histone modifications. For example, SUZ12, a member of the PRC2 complex, binds DNA in a sequence-specific manner and NRCF and ZBTB33 recruit histone deacetylase. The ENCODE project demonstrated a specific histone modification profile around binding sites of many TFs².

Since the same lysine residue cannot be both methylated and acetylated, the presence of H3K27ac is negatively correlated with the presence of H3K27me3³. Although there are several pieces of evidence showing that H3K27me3 is established at least partially in a sequence-specific manner, it is unclear if H3K27ac — an antagonistic activator to a repressive H3K27me3 — shares any sequence specific patterns. A computational approach⁴ predicted some TFBS to be linked to H3K27ac, yet the results were inconsistent and dependent on the background set. In this work, we perform a direct experiment to test whether a specific genomic sequence is capable of recovering H3K27ac.

Methods

Cas9 target selection and plasmid design

We selected a site for Cas9 in the intergenic region of human beta-globin locus using services **CCTop**⁵ (hg38) and **Off-Spotter**⁶ (hg38). The chosen targeting sequence was CTTGTCCCTGCA GGGTATTA. Then we designed targeting oligonucleotides (Table 2) for pSpCas9(BB)-2A-Puro (pX459 plasmid (Addgene), a gift from Egor Prokhorchouck, E.P.). Oligonucleotides Bgl1Cas_F and Bgl1Cas_R were diluted to 10 μ M each with annealing buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA), heated for 5 minutes to 95°C and then slowly cooled down to room temperature. 500ng of pX459 was digested with 5U of BpiI (Thermo Scientific) for 1h at 37°C, heat inactivated, run on 1.5% agarose, linearized plasmid was cut out and purified with QIAquick Gel extraction kit (Qiagen, cat. no. 28706) and ligated with Bgl1Cas_F/Bgl1Cas_R duplex. The ligation product was transformed into *E. coli* Top-10 cells, one of the clones was chosen and the insert was confirmed by Sanger sequencing. Below we refer to this plasmid as pX459-b1.

Plasmids design

Selection cassette design (HSV thymidine kinase, T2A peptide, NeoR in one frame) was performed with Ugene⁷. HSVtk and

NeoR sequences were PCR amplified with Phusion polymerase (Thermo Scientific) from pBS246-neo/Tk (a gift from E.P., construction of this plasmid is described elsewhere⁸) with primers HSVtk_F, HSVtk_R, and G418_F, G418_R, correspondingly. T2A peptide coding sequence, corresponding to amino acid sequence GSGEGRGSLTTCGDVEENPGP, was synthesized by hybridization of oligonucleotides T2A(+) and T2A(-) in annealing buffer and consequent treatment of hybridized duplex with T4 polymerase (Thermo Scientific) and gel purified. The NeoR fragment was digested with XbaI (Thermo Scientific), the T2A fragment was digested with NheI (Thermo Scientific), then these fragments were ligated and the fragment of expected size (881bp) was gel purified. This T2A-NeoR fragment was digested with XhoI (Thermo Scientific), HSVtk fragment was digested with SalI (Thermo Scientific), fragments were ligated and the fragment of predicted length was gel purified again. This HSVtk-T2A-NeoR fragment was double digested with EcoRI (Thermo Scientific) and BshTI (Thermo Scientific) and ligated with pX459-b1 double-digested with EcoRI and BshTI. This step gave the plasmid with full-length selection cassette. Below we refer to this plasmid as pHSVtk-T2A-NeoR. Then the pHSVtk-T2A-NeoR plasmid was double-digested with XbaI and Acc65I (Thermo Scientific), gel purified and ligated with hybridized AdaptUp(+)-AdaptUp(-) duplex. A successful insert was verified by digestion of newly introduced restriction sites: SalI and BamHI. We refer to this plasmid as pAdaptUp-HSVtk-T2A-NeoR. This plasmid was digested with NotI (Thermo Scientific), treated with FastAP (Thermo Scientific), gel purified and ligated with AdaptDown(+)-AdaptDown(-) duplex to introduce a 2xBpiI site that generates half-sites for BclI and XhoI after cleavage. The resulting plasmid was transformed into *E. coli* Top-10 cells. We refer to this plasmid as pAdaptUp-HSVtk-T2A-NeoR-AdaptDown. To obtain LoxP-flanked sequences, we used a pBK-CMV-derived plasmid with a modified multiple cloning site containing restriction sites for BclI, NheI, and XhoI. This plasmid was transformed into *E. coli* JM110 to eradicate Dam methylation, double digested with BclI and NheI, dephosphorylated, gel purified and ligated with LoxP(+)-LoxP(-) duplex, treated with T4 PNK. This plasmid was transformed into *E. coli* Top10 cells and referred to as pLoxP. To obtain homology regions that flank Cas9 cleavage site we PCR amplified them from Caki1 gDNA, using primer pairs Bgl1Up_F and Bgl1Up_R for an upstream fragment and Bgl1 Down_F and Bgl1Down_R for a downstream fragment. These PCR fragments were double digested with NheI and XhoI and ligated with pLoxP, double digested by the same sites and dephosphorylated. Ligation products were transformed into *E. coli* JM110, and purified plasmids were digested with BclI and XhoI to yield upstream and downstream fragments of DNA bearing LoxP site on its end. The upstream fragment was ligated with pAdaptUp-HSVtk-T2A-NeoR-AdaptDown double digested with SalI and BamHI to give pUp1L-HSVTK-T2A-NeoR-AdaptDown. The downstream fragment was ligated with pUp1L-HSVTK-T2A-NeoR-AdaptDown treated with BpiI and dephosphorylated. We refer to this plasmid as pUp1L-HSVTK-T2A-NeoR-L1Down.

Stably transfected cell line generation

We co-transfected Caki1 cells with the plasmids pX459-b1 and pUp1L-HSVTK-T2A-NeoR-L1Down using Lipofectamine 3000

(Thermo Scientific) in 2cm² wells following manufacturers instructions. After one week of Puromycin (3 µg/ml) selection, cells were split into 96-well plates and selected with 1mg/ml G418 for two weeks. Cells from successfully growing clones were split into two equal aliquots, one for growth and another for genomic DNA isolation. Clones were checked for the presence of the insert with a primer pair BGL1pcr_F - BGL1pcr_R surrounding the insert. One clone (referred to as Caki1-GcvS-G418R) with a homozygous insert was selected for the further experiments.

Construction of a recombination target

Plasmid pBK-CMV was digested with SacI and HpaI, blunted with T4 polymerase, self-ligated, transformed into *E. coli* Top10 and called pBCK-CMVdHpaI-SacI. LoxP(+) oligo was PNK treated and annealed with LoxP(-) to form a duplex. Bait(+) and Bait(-) oligos were annealed and ligated with hemi-phosphorylated LoxP duplex. Ligation products were resolved on 3% agarose gel, the longest fragment was excised and purified with QIAquick gel extraction kit, treated with T4 PNK and ligated to NheI treated and dephosphorylated pBCK-CMVdHpaI-SacI. The ligation product was transformed into *E. coli* Top10, we refer to the resulting plasmid as pIL-bait-L1. Ten sequences chosen to be inserted (Table 1) were PCR amplified from genomic DNA of Caki1 cells with primers (Table 2, rows 25 to 44). Amplicons were gel purified, diluted to a concentration of 100nM, treated with T4 PNK and ligated with a plasmid pIL-bait-L1 treated with Ecl136II and dephosphorylated. Then the library of ligation products was transformed into *E. coli* Top10, the plasmid library was purified using a plasmid mini kit (Evrogen, cat. no. BC021) and co-transfected with pBS598 EF1alpha-EGFPcre (Addgene) to Caki1-GcvS-G418R cells for Cre-mediated recombination exchange of the insert and the HSVTK-T2A-NeoR cassette. This step was performed in two independent replicates. After 3-day growth Ganciclovir (2 µM) was added to eliminate cells that did not undergo recombination. After selection for 10–14 days survived cells were grown to a

subconfluent monolayer in 10cm dishes and then ChIP on H3K27Ac was performed.

Chromatin immunoprecipitation (ChIP)

ChIP was performed according to [Abcam X-chip protocol](#) with the following modification: we increased the number of washes in High Salt buffer from one to three times. Sonication was performed in PCR tubes (SSIBio, cat. no. 3245-00) on ice using Sonics Vibra-Cell VCX 130 with an eight-element probe (cat. no. 630-0602). Sonication setup was: 10s pulse, 20s pause, 75% power, total sonication time is 30min. We used 2 µg of Anti-H3K27Ac antibody (Abcam, ab4729) per ChIP. Biological replicates of targeted inserts of DNA sequences were processed independently. Control ChIP was performed under the same conditions with Caki1 cells.

Nested PCR

After ChIP, the DNA was pooled and 1ng of DNA was PCR amplified with a primer pair Bait-seq_F – Bait-seq_R. At this step we amplified all DNA sequences inserted between this primer pair. Then 10ng of PCR amplicons from the first step was used as a template for a PCR with a primers (Table 2, rows 25 to 44), each pair in an individual reaction tube.

Results and discussion

None of the GC- and CpG rich promoter regions, that were acetylated in their original genomic loci (rows 1–3 in Table 1 and lanes 1–3 in Figure 1) recovered H3K27ac after relocation to a foreign genomic context in the beta globin locus, suggesting that H3K27ac may not depend directly on such features. An alternative explanation is that some of the CpG dinucleotides became methylated in a foreign genomic environment preventing chromatin from gaining H3K27ac - an active chromatin mark. Surprisingly, two extremely GC-rich but CpG poor (and, therefore, unmethylated) sequences (rows 5, 6 in Table 1 and lanes 5, 7 in

Table 1. Sequences inserted to foreign genomic context and their properties.

Sample number	Genome location, hg38	Gene name	Length, bp	GC, %	CpGs	H3K27Ac, native location	H3K27ac, foreign location	H3K27Me3, native location
1	chr1:62435851-62436130	USP1	280	59.6	15	yes	no	no
2	chr1:77979434-77979693	FUBP1	260	58	27	yes	no	no
3	chr1:86913854-86914119	HS2ST1	266	56.4	16	yes	no	no
4	chr20:63473349-63473556	-	208	65	1	no	no	yes
5	chr12:49033153-49033412	-	260	66.6	1	no	yes	no
6	chr17:81127213-81127450	-	238	67.7	2	no	yes	yes
7	chr2:227381301-227381590	-	290	26.6	1	no	no	no
8	chr1:108505575-108505854	-	280	37.5	0	no	no	no
9	chr19:50816997-50817284	-	288	51.7	0	yes (only GM12878)	no	yes (in some cells)
10	chrX:48988617-48988801	-	185	52	1	no	no	no

Table 2. Oligonucleotides used in this article.

Number	Name	Sequence
1	Bgl1Cas_F	CACCGCTTGTCCCTGCAGGGTATTA
2	Bgl1Cas_R	AAACTAATACCCTGCAGGGACAAGC
3	HSVtk_F	AATTACCGGTATGGCTTCGTACCCCTGCC
4	HSVtk_R	AATTGTCGACGTTAGCCTCCCCATCTCC
5	T2A(+)	AAAGCTCGAGGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTG
6	T2A(-)	ATTCGCTAGCTGGGCCAGGATTCTCCTCGACGTCACCGCATGTTAGCAGAC
7	G418R_F	ATTATCTAGAATTGAACAAGATGGATTGCACG
8	G418R_R	TAATGAATTCTCAGAAGAAGCTCGTCAAGAAGG
9	AdaptUp(+)	CTAGCAGTCGACTTAAGGATCCAT
10	AdaptUp(-)	GTACATGGATCCTTAAGTCGACTG
11	AdaptDown(+)	GGCCGCGATCCTGTCTTCAAGAAGACCTTCGAGC
12	AdaptDown(-)	GGCCGCTCGAAGGTCTTCTTGAAGACAGGATCGC
13	LoxP(+)	CTAGCATAACTTCGTATAATGTATGCTATACGAAGTTATT
14	LoxP(-)	GATCAATAACTTCGTATAGCATAATTATACGAAGTTATG
15	Bgl1Up_F	TCCACTCGAGACCTGGAAACCCATGTCG
16	Bgl1Up_R	GAATGCTAGCTCTGCGTTACACTCTAGTCACAC
17	Bgl1Down_F	CATCGCTAGCCCAGGCATACCAGGCAAATAAG
18	Bgl1_Down_R	TGCTCTCGAGAAATTGACACCATGGCCAC
19	BGL1pcr_F	TGCTGCAGATACCATCATCC
20	BGL1pcr_R	GTAGAATAGACCTGCACCTGCT
21	Bait(+)	GATCACACTCCTCTGAAGTGAGAGAGCTCCACTAGGACACCTTCTGGT
22	Bait(-)	GATCACCAGAAGGTGTCTAGTGAGCTCTCTCACTTCAGAGGAGTGT
23	Bait-seq_F	CACACTCCTCTGAAGTGAGAGAG
24	Bait-seq_R	CCAGAAGGTGTCCTAGTGAG
25	01_Hi_Ac_FUBP1_F	CCGCTATGGGTCATTCTCGG
26	01_Hi_Ac_FUBP1_R	ATTCAACACCCACCCTCGTG
27	02_Hi_Ac_USP1_F	GGCTGAACGTATCCTGCTTAAC
28	02_Hi_Ac_USP1_R	GCGAATGCTGTAGAGCTGAG
29	03_Hi_Ac_HS2ST1_F	GGGGAAGGAGAAGGATCAACC
30	03_Hi_Ac_HS2ST1_R	TGCGGAAAATGGTAGCGATG
31	04_No_Ac_GC66_F	GGCAGTAGGACCCCCTAGAC
32	04_No_Ac_GC66_R	GCATAACTAAGTCCAGGCCCC
33	05_No_Ac_GC67_F	CTGTGTCCCATAAGGCCCTG
34	05_No_Ac_GC67_R	ACAAACCAAGCTCTGGGTCC
35	06_No_Ac_GC68_F	CCCCCATGGCAAAGCAAG
36	06_No_Ac_GC68_R	CCCACCACCTCCAGAGGAG
37	07_No_Ac_GC28_F	CTATCAACACAACACACAAAACACC
38	07_No_Ac_GC28_R	AGGTAAGTAACTACCACCTCCCAC
39	08_No_Ac_GC39_F	AAATCCACTTTTCATCACCACCTC
40	08_No_Ac_GC39_R	TTTAAGAGAAATGGTTAATGTGGGG
41	09_No_Ac_GC51_F	TGAGGTGATTCTGTGGGGTAAAC
42	09_No_Ac_GC51_R	ACAACCCCAAAGTCTGGCAC
43	10_No_Ac_GC53_F	TTCAGACACACTCATCTCCCAC
44	10_No_Ac_GC53_R	GGTAACTGGGTCTGAAGATTCC

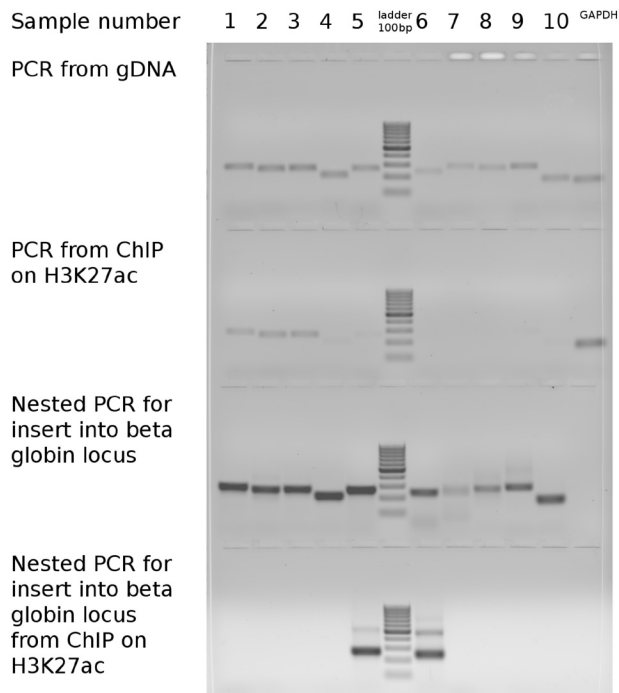


Figure 1. Amplification of chosen sequences in native and foreign genomic context; H3K27ac ChIP-PCR in native and foreign genomic context.

Figure 1) gained H3K27ac in the foreign environment, while in their native environment (in their original genomic location) they had no H3K27ac. Sequences 5 and 6 are located far from promoters of known genes. The sequence 5 contains a lowly expressed CAGE cluster, representing a weak alternative promoter of the KMT2D gene (ZENBU). Therefore, we conclude that the gain of H3K27ac in these regions is unlikely to be explained by transcriptional activity. The lack of H3K27ac in the sequence 6 in

the native location may be due to the presence of the antagonistic mark H3K27me3 in Caki1 cells, which is lost after translocation to a foreign environment. Yet, for the sequence 5, H3K27me3 is absent in all cell types reported in ENCODE. Our results suggest that in contrast to H2K27me3, H3K27ac gain is unlikely affected by the CpG content of the underlying DNA sequence, while extremely high GC-content might contribute to the gain of the H3K27ac.

Dataset 1. Gel image of amplification of target sequences from native genomic loci compared with amplification from beta-globin locus with or without ChIP on H3K27ac

<http://dx.doi.org/10.5256/f1000research.13441.d192774>

Data availability

Dataset 1: Gel image of amplification of target sequences from native genomic loci compared with amplification from beta-globin locus with or without ChIP on H3K27ac. DOI, [10.5256/f1000research.13441.d192774](https://doi.org/10.5256/f1000research.13441.d192774)⁹

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by the Russian Science Foundation [15-14-30002].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We are thankful to Egor Prokhorchouck for providing access to the laboratory equipment, cell lines, plasmid library and help with the experimental setup.

References

- Jermann P, Hoerner L, Burger L, *et al.*: Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc Natl Acad Sci U S A.* 2014; **111**(33): E3415–3421. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hoffman MM, Ernst J, Wilder SP, *et al.*: Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* 2013; **41**(2): 827–841. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tie F, Banerjee R, Stratton CA, *et al.*: CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development.* 2009; **136**(18): 3131–3141. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Whitaker JW, Chen Z, Wang W: Predicting the human epigenome from DNA motifs. *Nat Methods.* 2015; **12**(3): 265–272, 7 p following 272. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Stemmer M, Thumberger T, Del Sol Keyer M, *et al.*: CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One.* 2015; **10**(4): e0124633. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Pliatsika V, Rigoutsos I: “Off-Spotter”: very fast and exhaustive enumeration of genomic lookalikes for designing CRISPR/Cas guide RNAs. *Biol Direct.* 2015; **10**: 4. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Okonechnikov K, Golosova O, Fursov M, *et al.*: Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 2012; **28**(8): 1166–1167. [PubMed Abstract](#) | [Publisher Full Text](#)
- Prokhortchouk A, Sansom O, Selfridge J, *et al.*: Kaiso-deficient mice show resistance to intestinal cancer. *Mol Cell Biol.* 2006; **26**(1): 199–208. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Zubritskiy A, Medvedeva YA: Dataset 1 in: DNA sequence features in the establishing of H3K27ac. *F1000Research.* 2018. [Data Source](#)

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 04 June 2018

doi:10.5256/f1000research.14594.r34371



Oleg Gusev 

RIKEN, Yokohama City, Kanagawa, Japan

The study is focused on question of relation of primary genomic sequences and H3K27 acetylation preferences. Original method, based on the translocation of the different CG% and CpG -based sequences to the same genomic locus, followed by H3K27ac analysis was employed. The results suggest that that the context is more important than GC or CpG properties of the genomic sequence itself. The results of certain interest for the wide pool of researcher working with epigenetic aspects. At the same time, I believe that involvement of more techniques, such as Chip-Seq would be suggested to confirm the results. Also, it is not clear from the text how many replicates were taken and how general is the conclusion, if, for example, different genomic loci (or different genome, for example extremely GC or AT-rich one) will be used.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 12 Jul 2018

Anatoliy Zubritskiy, Research Center of Biotechnology RAS, Russian Federation

Q: The study is focused on question of relation of primary genomic sequences and H3K27 acetylation preferences. Original method, based on the translocation of the different CG% and CpG -based sequences to the same genomic locus, followed by H3K27ac analysis was employed. The results suggest that the context is more important than GC or CpG properties of the genomic sequence itself. The results of certain interest for the wide pool of researcher working with epigenetic aspects.

A: We thank the reviewer for a careful reading of the manuscript and useful comments. Although we have to disagree that our results show the context to be more important than GC or CpG content of the sequence. In fact, our results suggest that GC content may contribute to the establishing of H3K27ac. To avoid further confusions we modified the abstract and the results accordingly to make this statement more clear.

“Our results suggest that in contrast to H2K27me3, H3K27ac gain is unlikely affected by the CpG content of the underlying DNA sequence, while extremely high GC-content might contribute to the gain of the H3K27ac”

Q: At the same time, I believe that involvement of more techniques, such as Chip-Seq would be suggested to confirm the results.

A: We agree with the reviewer that a ChIP-seq may provide a more unbiased approach. Yet, a regular ChIP-seq experiment after the insertion seems to produce a lot of waste since only one region is expected to be changed. On the other hand a modified version of ChIP-seq experiment with multiple insertions into the same location may indeed be more unbiased and will hopefully increase the number of tested sequences by the order of magnitude making the conclusions more solid. We are working on this protocol now but we believe that this much more complicated experimental setup is beyond the scope of the current pilot paper.

Q: Also, it is not clear from the text how many replicates were taken and how general is the conclusion, if, for example, different genomic loci (or different genome, for example extremely GC or AT-rich one) will be used.

A: The insertion of 10 target sequences was performed done in two independent biological replicates, then every replicateate was CHI'P'd and resulting DNA was pooled before nested PCR step. To avoid any further confusions the corresponding text was added to the Methods.

Section “Construction of a recombination target”

“This step was performed in two independent replicates.” was added to the

Section “Chromatin Immunoprecipitation (ChIP)”

“Biological replicates of targeted insertion of DNA sequences were processed independently”

Section “Nested PCR”

“the DNA was pooled” was added to section “Nested PCR”

The beta-globin locus was chosen based on its lack of the chromatin modifications in the most of the adult cell types so the surrounding chromatin does not interfere with the inserted sequence and the gain of the modifications in the inserted fragment is determined by its sequence. We agree that additional loci lacking chromatin marks might be also tested, yet for now we focus on the modified ChIP-seq protocol to test many more inserted sequences.

At this stage we are not generalizing our conclusions to other genomes with the extreme AT- or GC-contents of the genome since apart of these changes these organisms may also have a quite different H3K27ac-attracting machinery.

Competing Interests: No competing interests were disclosed.

Referee Report 20 February 2018

doi:10.5256/f1000research.14594.r30746



Vasily N Aushev 

Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York City, NY, USA

This study aimed to test whether H3K27 acetylation depends on the properties of the target sequence itself, or rather defined by the surrounding genomic context. To assess this question, authors translocated sequences with different CG% and CpG content, to the same beta globin locus, and then tested their H3K27ac in the new location.

As three tested highGC-highCpG sequences lost their H3K27ac in the new location, authors conclude that the context is more important than GC or CpG properties of the sequence itself. At the same time, two other sequences actually acquire H3K27ac in the same new location, and only one of those can be possibly explained by the antagonistic H3K27me3 mark.

Authors describe their methods with all the details which should be sufficient for reproducibility of their main results. I had, however, some minor technical questions regarding their plasmids construction procedure. For example: 1) From the sequence of T2A oligos they provide, peptide sequence seems to be PSPLPSPCLLTCGDVEENPGP, not GSGEGRGSLTTCGDVEENPGP. The latter one would be present in the original pX459 plasmid, but not in the constructed one. 2) To my understanding, annealing of T2A(+) and T2A(-) oligos will produce a hybrid with only 18bp double-stranded overlap - for further digestion with NheI, its single-stranded ends need to be built up with some polymerase; I don't see this step in the Methods. 3) I don't see XhoI site in the T2A-NeoR fragment, and don't see any other potential site compatible with Sall of HSVtk fragment. In addition, it is recommended that all the source plasmids used should be properly referred - either by their original publication, or by the sequence reference in any available repository (Genbank, Addgene, etc); referring as "pHSVTK-Neo (a gift from E.P.)" is not sufficient if the plasmid's sequence cannot be explicitly found by that name. Finally, authors should double-check their text for typos in the names of all constructs and genes (for ex., GAPGH), be sure to use all the titles exactly as they are originally defined (for ex., PX459 instead of pX-459, etc). All these issues, however, are minor and do not distort the main content of the paper.

From the technical point of view, the only question I had is about the final readout for H3K27ac: authors perform nested PCR followed by visualization of the PCR products on the gel. I am wondering this method was chosen and not qPCR (as more quantitative) or ChIP-seq (as more unbiased).

One of my main concerns is about the reproducibility of the findings: authors do not report any biological or technical replicates which makes the results potentially questionable.

Talking about the interpretation of the results, I am not sure how strong can be the conclusion based on only 10 tested sequences, and even more so, only one genomic context. Even with this limited number of tests, some results cannot be easily explained (for example, acquired H3K27ac for sample #5), so in future it would be necessary to expand the panel of both sequences and contexts.

Besides above-mentioned questions (hopefully authors can address at least some of them in the revised version), I think the paper is clearly of scientific interest and adds important piece of knowledge to its field.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Referee Expertise: Cancer biomarkers, oncogenesis, signal transduction, miRNA regulation

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 12 Jul 2018

Anatoliy Zubritskiy, Research Center of Biotechnology RAS, Russian Federation

We would like to thank the reviewer for the thorough reading of our paper and very useful comments.

Q: From the sequence of T2A oligos they provide, peptide sequence seems to be PSPLPSPCLLTCGDVEENPGP, not GSGEGRGSLTTCGDVEENPGP. The latter one would be present in the original pX459 plasmid, but not in the constructed one.

A: We have to politely disagree with the reviewer on this point. According to <https://eu.idtdna.com/calc/analyzer> in hetero-dimer mode, the annealed T2A(+)/T2A(-) duplex with minimal dG value looks as follows:

AAAGCTCGAGGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTG

|||||||

CAGACGATTGTACGCCACTGCAGCTCCTCTTAGGACCGGGTCGATCGCTTA

For a double-stranded DNA synthesized with the T4 polymerase, the (+) strand looks as follows:

AAAGCTCGAG

GGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTGAGGAGAATCCTGGCCCA
GCTAGCGAAT

This sequence contains two restriction sites (XhoI and NheI). The sequence between restriction sites is translated as GSGEGRGSLTCDVEENPGP in 5'→3' direction, frame 1 (<https://web.expasy.org/cgi-bin/translate/dna2aa.cgi>).

Q: To my understanding, annealing of T2A(+) and T2A(-) oligos will produce a hybrid with only 18bp double-stranded overlap - for further digestion with NheI, its single-stranded ends need to be built up with some polymerase; I don't see this step in the Methods.

A: We thank the reviewer for pointing out this mistake. The typo in Methods section is corrected and enzyme name is changed from T4 PNK to T4 Polymerase. We modified the Methods section and added the following text:

"T2A peptide coding sequence corresponding to amino acid sequence GSGEGRGSLTCDVEENPGP was synthesized by hybridization of oligonucleotides T2A(+) and T2A(-) in annealing buffer and consequent treatment of hybridized duplex with T4 polymerase (Thermo Scientific) followed by gel purification."

Q: I don't see XhoI site in the T2A-NeoR fragment, and don't see any other potential site compatible with Sall of HSVtk fragment.

A: The T2A oligonucleotide duplex contains an XhoI site (**CTCGAG**) in a 5' region

AAAGCTCGAG

GGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCA

and a primer HSVtk_R contains a Sall restriction site (**GTCGAC**) **AATTGTCGAC**

GTTAGCCTCCCCATCTCC compatible with the XhoI site.

Q: In addition, it is recommended that all the source plasmids used should be properly referred - either by their original publication, or by the sequence reference in any available repository (Genbank, Addgene, etc); referring as "pHSVTK-Neo (a gift from E.P.)" is not sufficient if the plasmid's sequence cannot be explicitly found by that name. Finally, authors should double-check their text for typos in the names of all constructs and genes (for ex., GAPGH), be sure to use all the titles exactly as they are originally defined (for ex., PX459 instead of pX-459, etc). All these issues, however, are minor and do not distort the main content of the paper.

A: We carefully checked all the names of the genes and plasmids as well as the rest of the text for typos. The plasmid name pHSVTK-Neo is internal; this plasmid is referred as pBS246-neo/tk in the following paper: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1317619/>. The reference describing a the plasmid pBS246-neo/tk is added to the main text:

"HSVtk and NeoR sequences were PCR amplified with Phusion polymerase (Thermo Scientific) from plasmid pBS246-neo/Tk (a gift from E.P.), construction of this plasmid is described elsewhere: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1317619/>"

Q: From the technical point of view, the only question I had is about the final readout for H3K27ac: authors perform nested PCR followed by visualization of the PCR products on the gel. I am wondering this method was chosen and not qPCR (as more quantitative) or ChIP-seq (as more unbiased).

A: For a pilot study we decided to focus on making sure that the insertion of the selected fragments into a foreign environment (beta-globin locus) did take place as a proof of principle therefore we have chosen to perform nested PCR. In this regards qPCR will not make the conclusions more reliable. Performing a regular ChIP-seq experiment after the insertion seems like a bit of overkill since histone mark only in one region is expected to be changed. On the other hand a modified version of ChIP-seq experiment with multiple insertions may indeed be more unbiased. Yet, we

believe that this much more complicated experimental setup is beyond the scope of the current paper.

Q: One of my main concerns is about the reproducibility of the findings: authors do not report any biological or technical replicates which makes the results potentially questionable.

A: The insertion of 10 target sequences was performed in two independent biological replicates, then every replicate was ChIP'd independently and resulting DNA was pooled before nested PCR step. To avoid any further confusions the corresponding text was added to the Methods:

Section "Construction of a recombination target"

"This step was performed in two independent replicates." was added to the

Section "Chromatin Immunoprecipitation (ChIP)"

"Biological replicates of targeted insertion of DNA sequences were processed independently"

Section "Nested PCR"

"the DNA was pooled"

Q: Talking about the interpretation of the results, I am not sure how strong can be the conclusion based on only 10 tested sequences, and even more so, only one genomic context. Even with this limited number of tests, some results cannot be easily explained (for example, acquired H3K27ac for sample #5), so in future it would be necessary to expand the panel of both sequences and contexts.

A: We agree with the reviewer that some of the results (seq #5) are puzzling although reproducible since we performed all the tests in two biological replicates. At this stage we decided to report the results of the pilot project and focus on the performing a more complex experiment based on ChIP-seq as mentioned above which hopefully will increase the number of tested sequences by the order of magnitude.

Q: Besides above-mentioned questions (hopefully authors can address at least some of them in the revised version), I think the paper is clearly of scientific interest and adds important piece of knowledge to its field.

A: Again, we are grateful to the reviewer for the high evaluation of our work, careful reading of the manuscript and useful comments.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research