Incomplete erasure of histone marks during epigenetic reprogramming in medaka early development

Authors/Affiliations

Hiroto S. Fukushima[1], Hiroyuki Takeda[1] and Ryohei Nakamura[1]

[1]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, 113-0033, Japan

Correspondence to: htakeda@bs.s.u-tokyo.ac.jp, rynakamura@bs.s.u-tokyo.ac.jp

Supplemental methods

Spike-in ChIP-seq library preparation and sequencing

ChIP experiment was performed as previously described (Fukushima et al. 2019) with modifications. Fertilized eggs were collected, dechorionated using hatching enzyme, and incubated at 28°C. Embryos at appropriate stages were transferred into ice-cold PBS in 1.5 mL or 2 mL tube, homogenized by pipetting with P1000 tip, centrifuged for 10 minutes at 4°C and the supernatant was removed for deyolking. For zebrafish fibroblast cell preparation, trypsinized cells were collected and washed with PBS. Subsequently, the deyolked medaka embryonic cell pellet or zebrafish fibroblast cell pellet was resuspended with ice-cold PNPP (PBS containing 20 mM sodium butyrate, 1 mM PMSF, and 1 × protease inhibitor), and cells were cross-linked by adding formaldehyde (1% volume per volume final) for 8 minutes at room temperature then quenched by adding glycine (125 mM final). After washing with ice-cold PNPP, cross-linked cells were stored at -80°C as a dry pellet. The cross-linked cell pellet was lysed with Lysis-ChIP buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 20 mM sodium butyrate, 1 mM PMSF, 1 × protease inhibitor) and sonicated in a microTUBE AFA Fiber Snap-Cap 6x16mm using Covaris S220 with optimized parameters (Peak Power: 105, Duty Factor: 4.0, cycles per burst: 200, duration: 750 seconds).

To measure the concentration of chromatin and to confirm successful fragmentation, 10 μL of sonicated chromatin solution was treated with RNase A for 1 hour and Proteinase K for 2 hours, and DNA was purified by phenol: chloroform: isoamyl alcohol method and ethanol purification. The DNA concentration was measured by Qubit dsDNA HS Assay Kit, and the average of two tube concentrations was used for subsequent spike-in ChIP. Size distribution of fragmented DNA was examined using Agilent 2100 Bioanalyzer and DNA High Sensitivity Kit (Supplemental Table S2).

30 μL of Dynabead Protein A and 3 μg antibody (Supplemental Table S1) were incubated with RIPA buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1 mM EDTA, 0.5mM EGTA, 1 % Triton X-100, 0.1 % SDS, 0.1 % sodium deoxycholate) overnight at 4°C. 30 ng of experimental chromatin (from medaka embryos) and 15 ng or 30 ng of reference chromatin (from zebrafish fibroblasts) were used for H3K27ac / H3K27me3 / H3K9me3 ChIP or H3K4 methylation ChIP, respectively. Chromatin solution was incubated overnight at 4°C with RIPA-ChIP buffer (RIPA buffer supplemented with 20 mM sodium butyrate, 1 mM PMSF, 1 × protease inhibitor). Samples were washed 3 times with ice-cold RIPA buffer and once with ice-cold ChIP-TE buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA), and reverse-crosslinked by incubating in a thermomixer (65°C, 800 rpm) overnight with Lysis / NaCl buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 300 mM NaCl). Samples were placed on a magnetic stand, and the supernatant was transferred into new tubes,

treated with RNase A at 37° C for 2 hours, Proteinase K at 55°C for 2 hours, and DNA was purified by phenol: chloroform: isoamyl alcohol method and ethanol purification. The DNA concentration was measured by Qubit dsDNA HS Assay Kit.

ChIP-seq libraries were prepared using KAPA Hyper Prep Kit from 1 ng ChIPed DNA. When the amount of ChIPed DNA was less than 1 ng, ChIP was repeated once more, and the two tubes were combined and used for library preparation. All ChIP-seq libraries were sequenced using the Illumina HiSeq1500 system.

ATAC-seq

ATAC-seq using A485-treated or DMSO-treated embryos were conducted as previously described (Nakamura et al. 2021) with modification. Four dechorionated embryos at the late-blastula stage were homogenized in ice-cold PBS, centrifuged at 500 g, 4°C for 5 minutes, and the supernatant was removed. After washing with ice-cold PBS, cells resuspended in 50 μL of ice-cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, and 0.1% Igepal CA-630), centrifuged for 10 min at 500g, and the supernatant was removed. 50 μL of transposition mix (including 25 μL of TD and 2.5 μL of TDE1 from Nextera Sample Preparation kit) was added and the tube was incubated at 37°C for 30 minutes, and the tagmented DNA was purified using a MinElute PCR purification kit. Subsequently, small DNA fragments were size-selected and amplified by two sequential rounds of PCR. First, nine-cycle PCR was performed using indexed primers from a Nextera Index kit and KAPA HiFi HotStart ReadyMix, and the amplified DNA was size-selected to a size of <500 bp using AMPure XP beads. Then, a second five-cycle PCR was performed using the same primer as the first PCR and purified by AMPure XP beads. ATAC-seq libraries were sequenced using the Illumina NovaSeq 6000 system.

Biological replicates

Immunofluorescence experiments and phenotype analyses were repeated at least twice. Two biological replicates were generated for spike-in ChIP-seq (except the A485 experiment), ATAC-seq and RNA-seq. Three biological replicates were generated for spike-in ChIP-seq of A485-treated embryos.

Statistical analysis

All box and whisker plots in this study indicate the median (center line), the first and third quartiles (bottom and top edge of the boxes, respectively), and 1.5 times the interquartile range (whiskers). Outliers are not plotted except for images of immunofluorescence staining.

For statistical testing of multiple samples, normality and equal variances were tested first

by the Shapiro-Wilk test and the Bartlett test, respectively, with p-value = 0.05 using Python library SciPy (Virtanen et al. 2020). The null hypothesis was not rejected in the test in Supplemental Fig. S11D, while rejected in other tests. Therefore, we performed a Tukey-Krammer test in Supplemental Fig. S11D and Dwass, Steel, Critchlow and Fligner all-pairs comparison test in other figures using the Python package scikit_posthocs (Terpilowski 2019) (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, respectively). We performed a two-sided Fisher's exact test only for the multiple comparison of proportion of embryos in Figure 4C, and p-values were normalized by the Holm method using a Python library statsmodels (statsmodels.stats.multitest.multipletests, method="holm") (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, respectively).

For statistical testing of two samples, normality and equal variances were tested first by the Shapiro-Wilk test and the F-test, respectively, with p-value = 0.05, and the null hypothesis was rejected. Therefore, we performed the Wilcoxon's rank-sum test in Supplemental Fig. S16B using SciPy (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, respectively).

Quantification of immunofluorescence staining results

Fiji (Schindelin et al. 2012) was used for immunofluorescence staining quantification. For measurements of chromatin signal intensity, DAPI-positive regions were chosen and the average of intensities inside DAPI-positive regions in a single embryo was measured. For measurements of background signal intensity, a region outside the nuclei was manually chosen, and signal intensity in that region was measured. Finally, the signal intensity fold change (DAPI-positive region / background) was calculated and used for subsequent statistical tests. In the boxplots, each dot indicates the average intensity of 1-3, 3-5, ~40, or ~100 cells in a single broad field slice image of 16-cell, 64-cell, late morula, or late blastula embryo, respectively. Interphase (I) and mitotic phase (M) embryos at the 16-cell stage and 64-cell stage are separately shown. Phases of cell cycle after the late morula stage are not shown because cells divide asynchronously and are mostly in interphase from the late morula stage.

We note that we tested whether the difference in numbers of nuclei affect the results of quantification or not using H3K27ac data at the late blastula stage. We randomly chose four nuclei (i.e. down-sampling) for average calculation in one slice and compared the boxplot with and without down-sampling. We tried this for twenty times, but we could not find large difference between the data with and without down-sampling (Supplemental Fig. S2G). Therefore, we used all nuclei in one slice to calculate average intensity as above.

Quantification of the fraction of mitotic cells

Numbers of mitotic cells in the immunofluorescence staining images were manually counted. The

numbers of total cells were counted during the process of quantification of immunofluorescence staining.

Spike-in ChIP-seq and ATAC-seq data processing

Low quality reads and adapter-derived sequences were trimmed by Trimmomatic (Bolger et al. 2014). Trimmed-reads were aligned to the medaka (HdrR) and zebrafish (Zv9) concatenated genome and medaka HdrR genome using BWA (Li and Durbin 2010) for ChIP-seq and ATAC-seq, respectively. Alignments with mapping quality smaller than 20 and PCR duplicates were removed by SAMtools (Li et al. 2009). Spike-in ChIP-seq reads aligned to medaka or zebrafish genome were separated for subsequent analysis. For ATAC-seq data, MACS2 (Zhang et al. 2008) was used to generate signals per million reads tracks with following options: --nomodel --extsize 200 --shift -100 -g 600000000 -q 0.01 -B -SPMR, and this value in each peak was used for analysis of chromatin accessibility. Mapping information are shown in Supplemental Table S4.

Previous ChIP-seq data processing

Previous ChIP-seq data (Nakamura et al. 2014) (Supplemental Table S3) were aligned to medaka HdrR genome, and other processing was performed as described for spike-in ChIP-seq data processing.

Previous ATAC-seq data processing

The previous ATAC-seq data (Nakamura et al. 2021) (Supplemental Table S3) were downloaded and processed as described above.

RPKM calculation

We calculated conventional RPKM (RPKMconv) using the following formula:

$$\mathrm{RPKMconv} = (N \times 10^9) / (T \times S),$$

where N is the number of reads mapped in each region, T is the total read number and S is the size (bp) of each region. Also, based on a previous study (Li et al. 2014), we calculated spike-in normalized RPKM (RPKMspike) using the following formula:

$$\mathrm{RPKMspike} = \mathrm{RPKMconv} \times \mathrm{Tc\_exp} \times \mathrm{Ti\_ref} / (\mathrm{Tc\_ref} \times \mathrm{Ti\_exp})$$

$$= (\mathrm{Nc\_exp} \times 10^9 \times \mathrm{Ti\_ref}) / (\mathrm{Tc\_ref} \times S \times \mathrm{Ti\_exp}),$$

where Nc_exp is number of ChIP reads of experimental chromatin mapped in each region, Tc_exp is number of total ChIP reads of experimental chromatin, Tc_ref is number of total ChIP reads of reference chromatin, Ti_exp is number of total input reads of experimental chromatin, and Ti_ref is number of total input reads of reference chromatin.

ChIP-seq correlation

To compare spike-in ChIP-seq with conventional ChIP-seq, or between two replicates, the RPKMconv or RPKMspike of each 5 kb window was calculated for the whole genome. Pearson's correlation between two replicates suggested high reproducibility, so we pooled two replicates for subsequent analysis.

ATAC-seq correlation

To compare ATAC-seq data with previous data or between two replicates, the RPKMconv of each 5 kb window was calculated for the whole genome. Pearson's correlation between two replicates suggested high reproducibility.

Removal of signal-artifact regions in ChIP-seq

To remove regions containing repeat sequences in medaka genome, signal-artifact regions in ChIP-seq were identified as follows: MACS2 (Zhang et al. 2008) peak calling of input data was performed using following parameters: --broad --nomodel --extsize 250 -g 600000000 -q 0.001 -B --SPMR --keep-dup all, called peaks within 500 bp distance were concatenated using bedtools merge (Quinlan and Hall 2010), and peaks called by more than half of input samples were included in the signal-artifact regions. The signal-artifact regions were excluded from the analysis.

ChIP-seq peak calling and generating tracks

MACS2 (Zhang et al. 2008) was used to call peaks and to generate signals per million reads tracks using following parameters: -g 600000000 -q 0.01 -B --SPMR --keep-dup all and -g 1300000000 -q 0.01 -B --SPMR --keep-dup all for experimental chromatin and reference chromatin, respectively. The peaks from the four stages (16-cell, 64-cell, late morula and late blastula) were concatenated if they are within 150 bp using bedtools merge (Quinlan and Hall 2010), and peaks that overlapped with signal-artifact regions were removed. The peaks in unanchored contigs were also removed, and the remaining peaks were used for subsequent analyses. To measure the background enrichment level, random peaks were generated by bedtools shuffle (Quinlan and Hall 2010) from the merged and survived peaks. For spike-in normalization of track data, the values of normalized pileup track data generated by MACS2 were further normalized as follows:

$$Vspike = Vconv \times Tc\_exp \times Ti\_ref / (Tc\_ref \times Ti\_exp),$$

where Vspike is track data value after spike-in normalization, Vconv is original track data value generated by MACS2 --SPMR, and other variables are explained above in the RPKM calculation section. Integrative Genome Viewer (Robinson et al. 2011) was used for visualization of track data

before and after spike-in normalization.

Global histone modification level calculation

Based on a previous paper (Li et al. 2014) with slight modification, the global histone modification level of experimental chromatin (G) was calculated using spike-in ChIP-seq data as follows:

$$G = (Tc\_exp \times Ti\_ref) / (Tc\_ref \times Ti\_exp),$$

where variables are explained above in the RPKM calculation section.

Peak annotation

All merged peaks that survived after signal-artifact regions removal were associated to nearby genes using Homer annotatePeaks.pl (Heinz et al. 2010). Specifically, peaks within 2.5 kb upstream to 1.5 kb downstream of the transcription start site (TSS) of nearby genes were regarded as peaks in the promoter, and default output of Homer annotatePeaks.pl was used for other annotations ("Exon", "Intron", "TTS (transcription termination site)", "Intergenic").

Clustering analysis of H3K27ac peak

We chose peaks within 2.5 kb upstream to 1.5 kb downstream of TSS of nearby genes and peaks in distal regions (at least 2 kb away from promoters) as promoter peaks and enhancer peaks, respectively. If more than two peaks were included in the same promoter, only the representative peak whose RPKMspike was highest was used for promoter peak analysis. RPKMspike value within each peak was used for subsequent clustering. To determine the optimal number of clusters for k-means clustering, we applied the elbow method and the Silhouette method, and we also checked the heatmaps from various numbers of clusters (see also Supplemental Figure S12, S14). Finally, we concluded that 5 and 6 is the best number of clusters for promoter and enhancer peak clustering, respectively.

Classification of H3K27me3 peaks

H3K27me3 peaks were divided into three groups using RPKMspike values of peaks in the 64-cell stage (“Erasure”: RPKMspike = 0, “Low retention”: $0 < RPKMspike \leq 1$, “Mild retention”: $1 < RPKMspike$). H3K27me3 peaks within 2.5 kb upstream to 1.5 kb downstream of TSS of nearby genes were considered as promoter peaks. If more than two peaks were included in a same promoter, a representative peak whose RPKMspike is highest was used.

Gene Ontology analysis

To find Gene Ontology terms enriched in promoter peaks, the genes downstream of H3K27ac or

H3K27me3 promoter peaks were listed. Gene Ontology enrichment analysis was performed by PANTHER (Mi et al. 2019). In Supplemental Figure S13C, Gene Ontology terms (biological process) which have FDR less than 10^{-5} and whose enrichment was more than 2.1 fold were listed. In Figure S19F, Gene Ontology terms (biological process) which have FDR less than 10^{-5} and whose enrichment was more than 3.3 fold were listed. Because many terms were found enriched in the "Mild retention" cluster, the top 20 terms sorted by FDRs were shown in Supplemental Figure S19F.

RNA-seq data processing

Previous RNA-seq data (Ichikawa et al. 2017; Nakamura et al. 2021) and RNA-seq data of A485-treated embryos were processed as previously described (Nakamura et al. 2021) (Supplemental Table S3). Low quality reads and adapter-derived sequences were trimmed by Trimmomatic (Bolger et al. 2014). The reads were aligned to medaka HdrR genome using STAR (Dobin et al. 2013), and alignments with mapping quality smaller than 20 were removed using SAMtools (Li et al. 2009). FPKM, RPKM, and adjusted p-value were generated by R library DESeq2 (Love et al. 2014). We regarded the genes whose adjusted p-value (padj) was less than 0.01 and whose log2 fold change (A485/DMSO) was more than 1 or less than -1 as significantly altered genes.

Analysis of maternal expression levels

RNAs in 2-cell-stage embryos are maternally derived. Therefore, we used FPKM at the 2-cell stage for analysis of maternal expression levels.

Whole genome bisulfite-seq data processing

The whole genome bisulfite-seq data from a previous study (Qu et al. 2012) (Supplemental Table S3) was downloaded from accession number SRA026693. Low quality reads and adapter-derived sequences were trimmed by Trimmomatic (Bolger et al. 2014). Trimmed reads alignment to medaka HdrR genome, deduplication and CpG methylation level calculation were done using bismark (Krueger and Andrews 2011). Only those CpG dinucleotides with coverage of ≥ 4 were regarded as valid calls and used for subsequent analyses.

Analysis of DNA methylation level around peaks

DNA methylation level averages in 100 bp window-divided genomic regions around peak center (within peak center ± 2 kb) were calculated and used for heatmap and line plot. If any valid CpG dinucleotides were not included in the 100 bp window-divided genomic region, the average methylation level of that region was accounted to be 100%.

Analysis of epigenetic changes induced by A485 treatment

Initially, we pooled two biological replicates of ChIP-seq and performed macs2 peak calling as described above. Here, we only chose peaks whose fold change of H3K27ac, H3K27me3 and H3K4me2 enrichment in macs2 peak calling output was higher than 5.8, 10.0 and 8.0, respectively. Those peaks from DMSO- or A485-treatment were next merged as described above. H3K27ac peaks that overlapped with H3K4me2 peaks were classified into “H3K4me2 nonsensitive” peaks ($Y \geq 0.8 \times X - 20$) and “H3K4me2 down” ($Y < 0.8 \times X - 20$) peaks (X and Y indicate H3K4me2 levels in DMSO-treated and A485-treated embryos, respectively, Fig. 6B). “H3K4me2 down” peaks overlapping with H3K27me3 peaks were regarded as “H3K4me2 down, H3K27me3 up”, because most of the peaks showed an increase in H3K27me3 (Supplemental Fig. S17B). We defined an H3K27ac peak as overlapping when its peak center is located inside a peak of another modification.

To compare the chromatin accessibility in H3K27ac peaks, ATAC-seq data was processed as described above. The peaks whose accessibility (output value of ATAC-seq MACS2) was higher than 2 at least in one sample and whose fold change of accessibility (log2(A485/DMSO)) was higher than 1.2 or lower than -1.2 were classified as significant changes.

References

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**: 576–589.

Ichikawa K, Tomioka S, Suzuki Y, Nakamura R, Doi K, Yoshimura J, Kumagai M, Inoue Y, Uchida Y, Irie N, et al. 2017. Centromere evolution and CpG methylation during vertebrate speciation. *Nat Commun* **8**: 1833.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–1572.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589–595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li X-Y, Harrison MM, Villalta JE, Kaplan T, Eisen MB. 2014. Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. *Elife* **3**.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* **47**: D419–D426.

Nakamura R, Motai Y, Kumagai M, Wike CL, Nishiyama H, Nakatani Y, Durand NC, Kondo K, Kondo T, Tsukahara T, et al. 2021. CTCF looping is established during gastrulation in medaka embryos. *Genome Res* **31**: 968–980.

Nakamura R, Tsukahara T, Qu W, Ichikawa K, Otsuka T, Ogoshi K, Saito TL, Matsushima K, Sugano S, Hashimoto S, et al. 2014. Large hypomethylated domains serve as strong repressive machinery for key developmental genes in vertebrates. *Development* **141**: 2568–2580.

Qu W, Hashimoto S, Shimada A, Nakatani Y, Ichikawa K, Saito TL, Ogoshi K, Matsushima K, Suzuki Y, Sugano S, et al. 2012. Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. *Genome Res* **22**: 1419–1425.

Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.

Terpilowski M. 2019. scikit-posthocs: Pairwise multiple comparison tests in Python. *J Open Source Softw* **4**: 1169.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**: 261–272.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.