## Supplemental Materials and Methods

### Construction and Sequencing of Bisulfite-treated DNA Libraries

Genomic DNA from medaka tissue cells was isolated using a DNeasy Kit (Qiagen). To obtain fragmented DNA of the desired size range (100 – 400 bp), 7 μg of the genomic DNA was sonicated for 3 min using a Covaris sonicator. The DNA fragments were treated with Klenow DNA polymerase to generate blunt end, and double–stranded DNA adaptors containing methylated cytosines were then ligated to the blunt-ended DNA fragments. The sequences of these methylation linkers were:

5′-C$^{m}$C$^{m}$AC$^{m}$C AC$^{m}$GC$^{m}$C$^{m}$ CC$^{m}$C$^{m}$GC$^{m}$ CCTC$^{m}$C$^{m}$ CC$^{m}$TC$^{m}$C ACGGG C$^{m}$AGTC$^{m}$ GGTGA T-3′,

5′-ATCAC CGACT GCCCG TGGAG GGGAG GGCGG GGGCG TGGTG GTT-3′,

5′-GACGC CCAAG AGAAC GAGGA AC$^{m}$C$^{m}$C$^{m}$G-3′, and

5′-TTTCG GGTTC CTCGT TCTCT TGGGC GTC-3′.

The DNA fragments were separated from the unligated adaptors by gel electrophoresis and then subjected to bisulfite treatment using an EpiTect Bisulfite Kit (Qiagen). The resulting bisulfite-treated DNA was used as a template for PCR with an initial denaturation step at 94°C for 5 min; 7–10 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and elongation at 72°C for 45 s; and a final extension step at 72°C for 5 min. The PCR primers used were:

converted linker-specific primer A (5′-AATGA TACGG CGACC ACCGA CAGGT TCAGA GTTCT ACAGT CCGAC CACTA CGCCT CCGCT TTCCT CTCTA T-3′), and

converted linker-specific primer B (5′-CAAGC AGAAG ACGGC ATACG ACTGC CCCGG GTTCC TCATT

CTCTT AAACA TC-3′),

which were designed to amplify only those DNA fragments carrying bisulfite-converted adaptor sequences at both ends. The PCR products were size-fractionated by 8% polyacrylamide gel electrophoresis, and those between 250 and 450 bp in length were recovered. The size-fractionated cDNAs were sequenced using an Illumina GAIIx genome analyzer according to the manufacturer's instructions. A validation experiment was performed according to the same procedure using genomic DNA from *Saccharomyces cerevisiae* S288C and its genome sequence (downloaded from http://genome.ucsc.edu/).

**Alignment of Bisulfite-treated Reads**

After removing low-quality reads (those containing five or more undetectable bases), we converted all cytosines in the remaining reads and in both the Watson and Crick strands of the reference genome to thymines before mapping them. The medaka genome and predicted gene sequences were downloaded from http://utgenome.org/medaka/. Using primary alignments between the converted reads and converted reference sequences, we created non-gapped Smith-Waterman alignments between the original sequences of these best hits. All possible methylation patterns were evaluated and only uniquely mapped reads were retained for further analyses. For 36-mer reads, no more than two mismatches [by Efficient Large-Scale Alignment of Nucleotide Databases (ELAND) mapping software] were allowed; for 76-mer reads, no more than three mismatches (by Burrows-Wheeler Aligner mapping software) were allowed. The average number of mismatches for the 76-mer reads was approximately one.

**Divergence Estimates and CpG Site Block Definition**

Divergence between the Hd-rR and HNI genomes has been detected through BLASTZ alignment analysis as described previously with all repetitive regions masked (Sasaki et al. 2009). A CpG site block was defined as an alignment block conserved between the two medaka genomes, consisting of at least one CpG site from either the Hd-rR or the HNI genome and its surrounding upstream or downstream ten bases (window size). Overlapping CpG site blocks were merged so that any two CpG sites separated by less than 20 bases were connected.

**Single- and Double-Nucleotide Substitution Rates in CpG Blocks of Human and Medaka Genomes**

To calculate the single-nucleotide substitution rates in the reference human genome (hg19), we used reference SNPs from the HapMap human CEU population (Utah residents with Northern and Western European ancestry; http://hapmap.ncbi.nlm.nih.gov/). Single-nucleotide substitution rates in the medaka genome were computed as the number of variant bases in the HNI genome divided by the total length of the corresponding regions in the Hd-rR genome. Gaps and undetectable nucleotide (N) variations were excluded. Because single-nucleotide mutations consisting of deaminated methylated cytosines in CpG dinucleotides are dominant, we excluded SNPs in CpG dinucleotides in this analysis.

Because the SNP rate in the human genome (~0.1%) was too low to determine the statistical significance of its dinucleotide substitution rates, we then calculated dinucleotide substitution rates for the medaka genome. We defined the dinucleotide substitution rates in the medaka genome as the number of dinucleotides with variant bases (excluding gaps or N) in the HNI genome, divided by the number of dinucleotides in the Hd-rR genome.

Dinucleotide transition and transversion rates were calculated as the number of dinucleotides with one transition or transversion base (excluding gaps or N) in the HNI genome, divided by the number of dinucleotides in the Hd-rR genome. Each dinucleotide had the same mutation rate as its reverse complementary dinucleotide.

**Methylation State of Cytosine Nucleotides and CpG Site Blocks in Medaka Genomes**

The level of methylation of a particular cytosine was estimated by dividing the number of reads reporting a cytosine (C) by the total number of reads reporting a C or T (thymine). Because methylation was assumed to occur simultaneously on both strands of medaka CpG sites, the methylation level for each CpG site was estimated using the sum of the occurrences in both strands to achieve stronger coverage. Only CpG sites that were covered 5 or more times were used to determine the average methylation level at a CpG site block**.** The overall average substitution rate for CpG site blocks was assessed among blocks including at least one CpG site that was covered by at least one bisulfite-sequencing read in both medaka strains. Blocks in which no methylation was detected because all CpG sites were mutated and blocks causing a high risk of mis-assembly, and these blocks were eliminated. The results are summarized in Table S3.

**Two proportion z-test**

Let $S_1$ and $S_2$ be disjointed sets of elements of unequal sizes, $n_1$ and $n_2$, respectively. Assume that an element is either positive or negative; e.g., a base in one genome is positive if it is mutated in another genome, and a CpG block is positive if it contains CGCG as a substring. Let $m_1$ and $m_2$ be the numbers of "positive" elements in $S_1$ and $S_2$, respectively. As an example, consider the disjointed sets of bases in hypo- and hypermethylated CpG site blocks, where positive elements are mutated bases. Suppose that the probabilities (means) of being a positive element in $S_1$ and $S_2$ are $p_1$ and $p_2$, respectively. We test the null hypothesis that the two probabilities are equal ($p_1 = p_2$) using the two-sample proportion *z*-test (George W. Snedecor 1989) (p.124) using a test statistic

$$z = \frac{\frac{m_1}{n_1} - \frac{m_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}, \text{ where } p = \frac{m_1+m_2}{n_1+n_2}.$$

We can calculate the *p*-value of the statistic $z$ based on the standard normal approximation as described below: Let $X_{1,i}$, $(i = 1, \dots, n_1)$ be independent and identically distributed random variables whose values are 1 if the $i$-th element in $S_1$ is positive and 0 if negative; namely, $P(X_{1,i} = 1) = p_1$ and $P(X_{1,i} = 0) = 1 - p_1$ for $i = 1, \dots, n_1$. Similarly, let $X_{2,i}$, $(i = 1, \dots, n_2)$ be independent and identically distributed random variables with $\mathrm{P}(X_{2,i} = 1) = p_2$ and $\mathrm{P}(X_{2,i} = 0) = 1 - p_2$ for $i = 1, \dots, n_2$. We denote by $\mathrm{N}(\mu, \sigma)$ the normal distribution with mean μ and standard deviation σ. From the central limit theorem, we know that the distribution of proportion of positive elements in $S_1$, $\sum_{i=1}^{n_1} X_{1,i}/n_1$, can be approximated by $\mathrm{N}(p_1, p_1(1-p_1)/n_1)$ and the distribution of $\sum_{i=1}^{n_2} X_{2,i}/n_2$ by $\mathrm{N}(p_2, p_2(1-p_2)/n_2)$ when $n_1$ and $n_2$ are sufficiently large. From the additivity of the normal distribution, we see that the difference of the two proportions, $\sum_{i=1}^{n_1} X_{1,i}/n_1 - \sum_{i=1}^{n_2} X_{2,i}/n_2$, approximates $\mathrm{N}(p_1 - p_2, p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2)$. By standardizing the difference, we obtain the

standard normal distribution for approximating the distribution of the test statistic $\frac{\frac{\sum_{i=1}^{n_1} X_{1,i}}{n_1} - \frac{\sum_{i=1}^{n_2} X_{2,i}}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$, where we substitute $p_1$ and $p_2$ by $p$ assuming the null hypothesis.

**Estimation of the Ancestor Sequence of Hd-rR and HNI**

A draft genome of an out-group medaka strain HSOK (*O. latipes* of South Korea) of Hd-rR and HNI was assembled using the SOAP de novo assembler (Li et al. 2010). The total number of contigs of the draft genome was 346,136, N50 contig size was 4378, and the maximum contig length was 117,159. Estimates of the ancestor sequence of Hd-rR and HNI were given by the following steps to calculate the evolution rate of hypo- and hypermethylated CpG site blocks longer than or equal to 100 bp. First, we performed pair-wise alignments of the Hd-rR and HNI sequences of the CpG site blocks to the HSOK genome using BLASTN; a CpG site block was retained for further analysis if both sequences in the block were mapped to the same position on the HSOK genome. Each base in the ancestor sequences was identified as the base most frequently observed at each position of the multiple alignment.

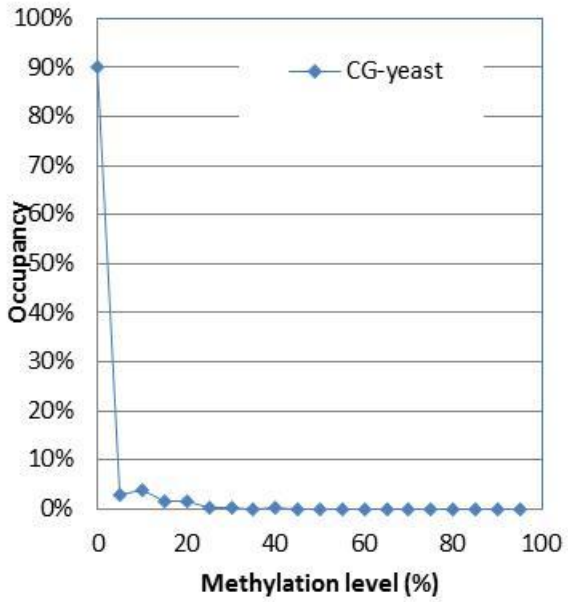
**Motif Detection in Hypo-/Hypermethylated CpG Site Blocks**

The AdaBoost algorithm was used to find conserved motifs in hypo-/hypermethylated CpG site blocks longer than or equal to 100 bp. All 3- to 6-bp conserved segments were considered to be candidate classifiers. Data sets (10,000 CpG site blocks each) were chosen from evolutionarily conserved hypo-/hypermethylated classes in blastulae, liver and testes samples.

## Supplemental Figures

**Figure S1.** Distribution of methylation in *Saccharomyces cerevisiae* S288C (control). (A) CpG cytosines. (B) Non-CpG cytosines. Cytosines with fewer than five reads were excluded because poorly covered CpG sites might reduce accuracy.

**A**




**B**

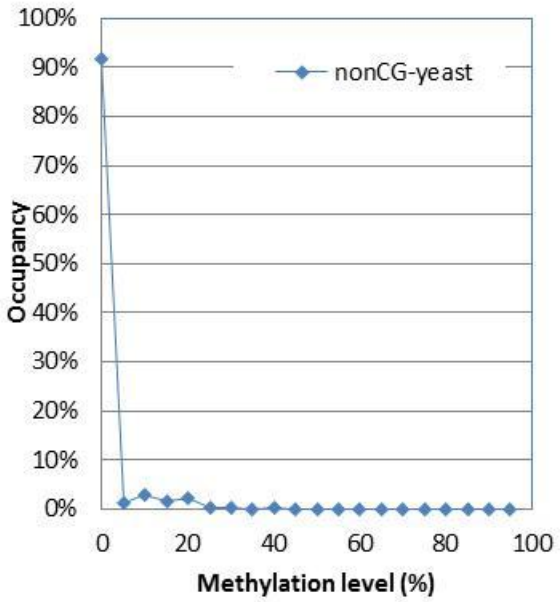
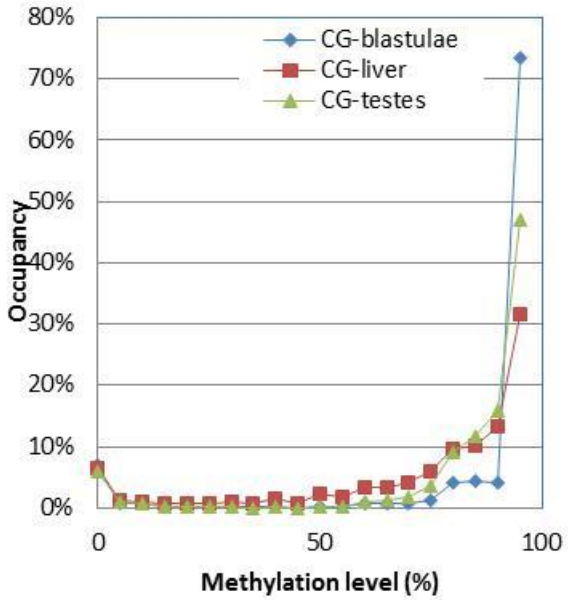
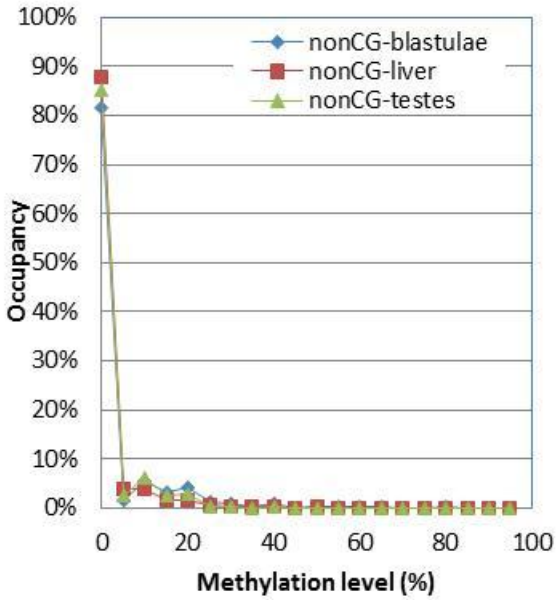


**Figure S2.** Distribution of methylation in Hd-rR and HNI. **(A)** CpG cytosines in Hd-rR. **(B)** Non-CpG cytosines in Hd-rR. **(C)** CpG cytosines in HNI. **(D)** Non-CpG cytosines in HNI. Cytosines covered with fewer than five reads were excluded because poorly covered CpG sites might reduce accuracy.

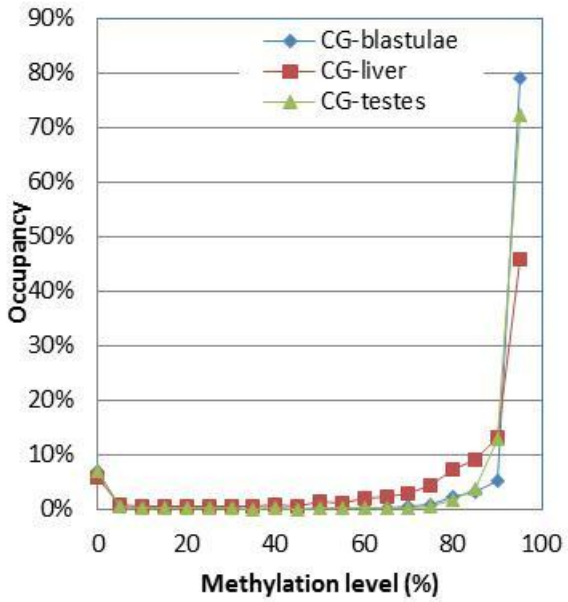
**A**




**B**




**C**




**D**

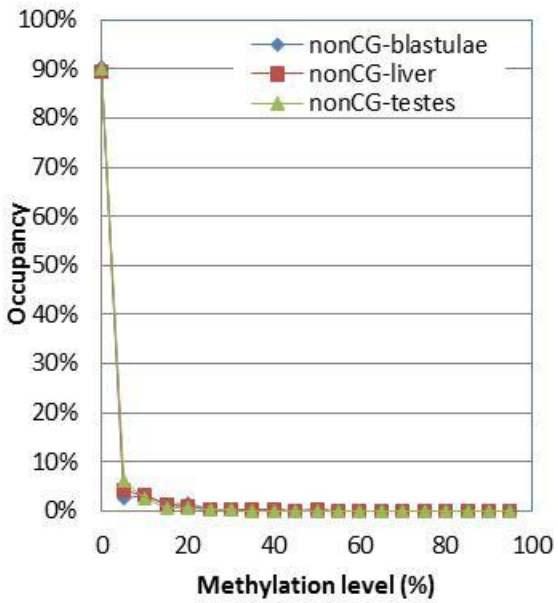
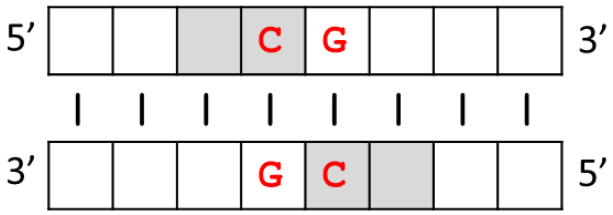
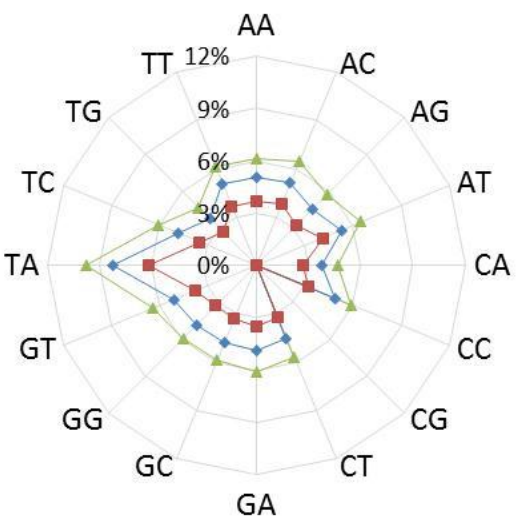


**Figure S3.** Methylation patterns and substitution rates in two medaka strains. **(A)** Substitutions of dinucleotides overlapping CpG by one base (shown in grey) were excluded. **(B-C)** Dinucleotide substitution rates in blastulae and testes excluding substitutions before CpGs as shown in (A).

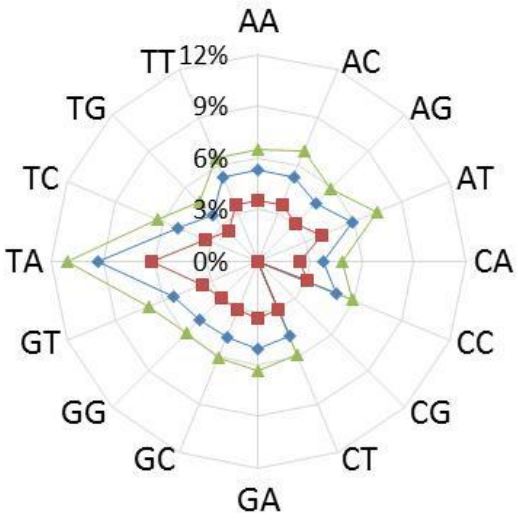
**A**




**B**




**C**



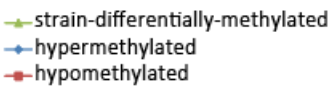
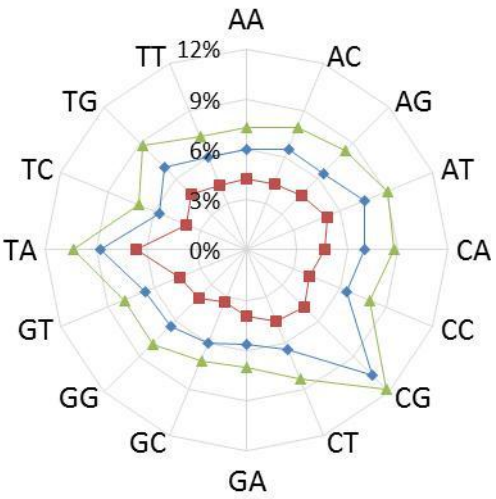
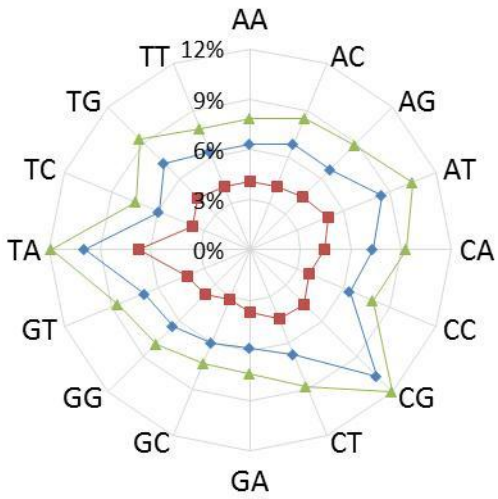



**Figure S4.** Methylation patterns and substitution rates in two medaka strains. CpG site blocks near TSSs (within 500 bp upstream or downstream) were excluded. **(A-B)** Dinucleotide substitution rates in blastulae and testes.

**A**




**B**



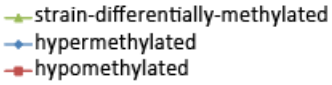
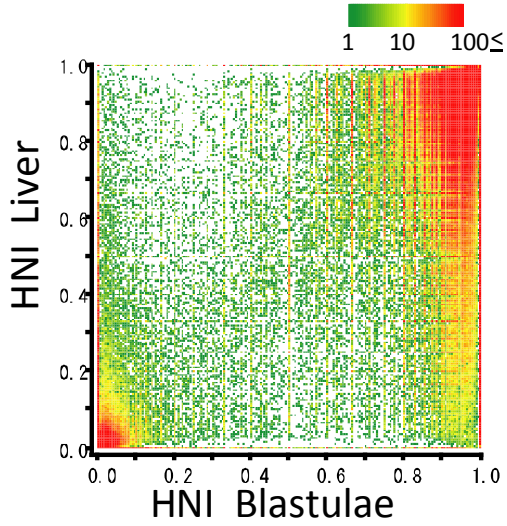



**Figure S5.** Comparison of HNI methylation patterns. **(A)** blastulae vs. liver, **(B)** Testes vs. liver.

**A**




**B**

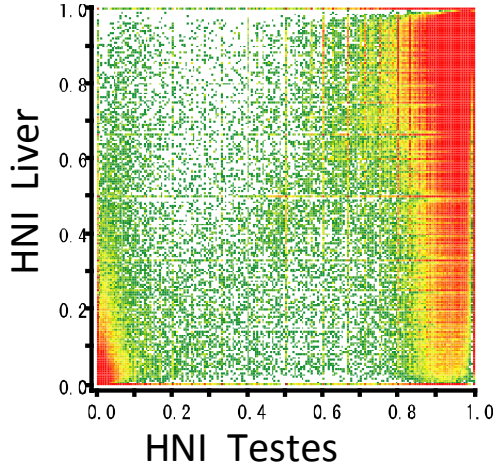
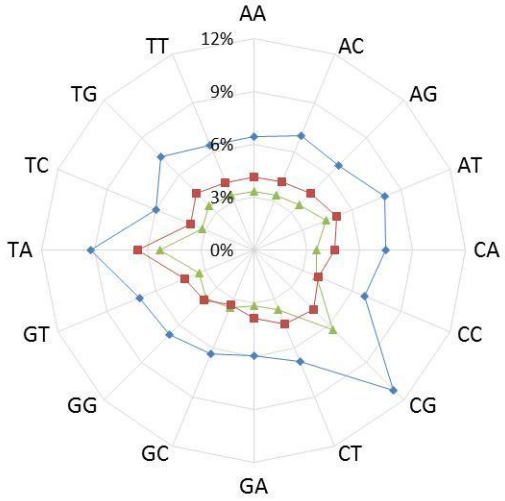
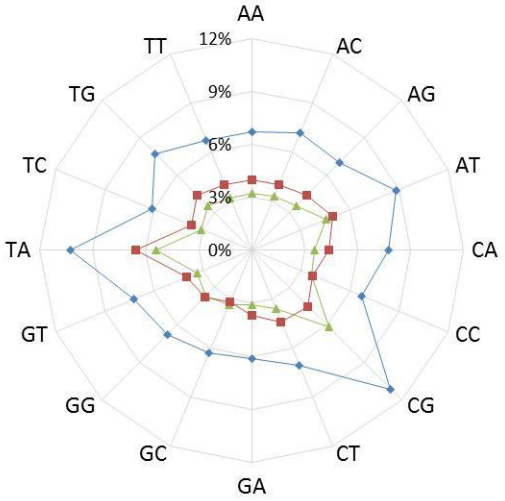


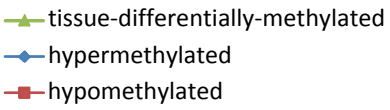
**Figure S6.** Methylation patterns and dinucleotide substitution rates in different cell lines. CpG site blocks present within 500 bp (upstream or downstream) of transcription start sites were excluded. **(A-B)** Dinucleotide substitution rates in CpG site blocks with various methylation states in blastulae vs. liver (**A**) and in testes vs. liver (**B**).

**A** (Blastulae vs Liver)




**B** (Testes vs Liver)



## Supplemental Tables

**Table S1.** Summary of bisulfite-sequencing and mapping in Hd-rR **(A)** and HNI **(B)** medaka genomes. *Genome coverage shows the average read depth under the assumption that the medaka genome is ~700 Mbp in length.

**A**

| | Number of raw reads | Number of uniquely mapped reads | Number of bases in the mapped reads | Genome coverage | Ratio of bases covered by ≥ 1 reads |
|---|---|---|---|---|---|
| blastulae | 278,558,149 | 53,952,009 | 4,100,352,684 | 585.5% | 76.8% |
| liver | 367,650,153 | 116,641,223 | 8,162,582,868 | 1165.5% | 85.5% |
| testes | 311,546,042 | 153,792,233 | 11,671,882,552 | 1666.5% | 86.4% |

**B**

| | Number of raw reads | Number of uniquely mapped reads | Number of bases in the mapped reads | Genome coverage | Ratio of bases covered by ≥ 1 reads |
|---|---|---|---|---|---|
| blastulae | 294,256,695 | 55,981,884 | 4,254,623,184 | 376.7% | 74.6% |
| liver | 306,831,850 | 95,934,117 | 7,290,992,892 | 645.5% | 81.4% |
| testes | 263,749,811 | 116,501,028 | 8,294,656,808 | 734.4% | 62.0% |

**Table S2.** Summary of bisulfite-sequencing and mapping for *Saccharomyces cerevisiae* S288C (control).

| | Number of raw reads | Number of uniquely mapped reads | Number of bases in the mapped reads | Genome coverage * | Ratio of bases covered by ≥ 1 reads |
|---|---|---|---|---|---|
| yeast | 18,198,425 | 588,276 | 44,708,976 | 183.8% | 66.1% |

**Table S3.** Number and length of CpG site blocks for various methylation states. **(A)** Number of CpG site blocks conserved in germline-like cells of Hd-rR and HNI. Exclusion of CpG site blocks within 500 bp (upstream or downstream) of transcription start sites is indicated by "not around TSS." **(B)** Lengths (bp) of CpG site blocks in germline-like cells. Nucleotides without gaps or undetectable nucleotide (N) lengths in Hd-rR are shown. **(C)** Comparisons of CpG site block methylation states between samples. **(D)** Comparisons of the combined lengths (bp) of the CpG site blocks in different samples. Nucleotides without gaps or undetectable nucleotide (N) lengths in Hd-rR are shown.

**A**. Number of CpG site blocks

| Methylation level conserved in Hd-rR and HNI | Blastulae | Testes | Blastulae (not around TSS) | Testes (not around TSS) |
|---|---|---|---|---|
| Hypermethylated | 2,068,068 | 2,062,244 | 1,996,300 | 1,985,697 |
| Hypomethylated | 116,505 | 106,983 | 84,421 | 77,001 |
| Strain-differentially-methylated | 35,887 | 30,039 | 33,295 | 27,754 |
| all blocks | 3,977,903 | 3,592,432 | 3,811,036 | 3,436,394 |

**B**. Length of CpG site blocks

| Methylation level conserved in Hd-rR and HNI | Blastulae | Testes | Blastulae (not around TSS) | Testes (not around TSS) |
|---|---|---|---|---|
| Hypermethylated | 76,888,469 | 79,913,082 | 74,071,594 | 76,865,285 |
| Hypomethylated | 5,936,587 | 5,621,367 | 4,175,689 | 3,935,026 |
| Strain-differentially-methylated | 1,356,113 | 1,244,558 | 1,256,856 | 1,152,318 |
| all blocks | 138,652,303 | 129,072,812 | 132,079,168 | 122,780,045 |

**C**. Number of CpG site blocks

| Methylation level conserved in Hd-rR and HNI | Blastulae vs. Liver | Testes vs. Liver | Blastulae vs. Liver (not around TSS) | Testes vs. Liver (not around TSS) |
|---|---|---|---|---|
| Hypermethylated | 1,185,258 | 1,215,534 | 1,143,555 | 1,170,007 |
| Hypomethylated | 104,894 | 96,784 | 74,944 | 68,812 |
| Strain-differentially-methylated | 69,079 | 59,285 | 65,637 | 56,069 |
| all blocks | 3,933,220 | 3,537,241 | 3,767,636 | 3,382,947 |

**D**. Length of CpG site blocks

| Methylation level conserved in Hd-rR and HNI | Blastulae | Testes | Blastulae vs. Liver (not around TSS) | Testes vs. Liver (not around TSS) |
|---|---|---|---|---|
| Hypermethylated | 48,562,440 | 51,591,871 | 46,758,079 | 49,601,547 |
| Hypomethylated | 5,500,344 | 5,223,772 | 3,817,064 | 3,612,601 |
| Strain-differentially-methylated | 2,213,649 | 1,929,369 | 2,095,109 | 1,817,688 |
| all blocks | 137,296,621 | 127,203,780 | 130,761,759 | 120,969,875 |

**Table S4.** The *z*- and *p*-values from the two-proportion *z*-test on the difference in SNP rates between hypo-, hyper-, and strain-differentially-methylated regions in human sperm and medaka blastulae grouped by the entire genome, intergenic regions, exons, and introns.

| | Human | | | | Medaka | | | |
|---|---|---|---|---|---|---|---|---|
| | Genome | Intergenic | Exons | Introns | Genome | Intergenic | Exons | Introns |
| hyper- vs hypo-methylated regions | | | | | | | | |
| z-value | 51.4 | 37.5 | 11.6 | 26.4 | 140.0 | 108.4 | 22.8 | 52.0 |
| p-value | <2.5E-567 | 1.3E-306 | 2.0E-30 | 4.7E-152 | <1E-2170 | <1E-2170 | 4.5E-114 | 5.2E-590 |
| strain-differentially- vs hyper- methylated regions | | | | | | | | |
| z-value | N/A | N/A | N/A | N/A | 44.2 | 36.1 | 7.3 | 20.5 |
| p-value | N/A | N/A | N/A | N/A | <7.3E-423 | 1.2E-283 | 8.0E-13 | 1.0E-92 |

**Table S5.** The *z*- and *p*-values from the two-proportion *z*-test of the difference in dinucleotide substitution rates between hypo- and hyper-methylated regions conserved in the two medaka strains. The maximum *p*-value in each category (genomes, intergenic regions exons, or introns) is indicated in bold type; all *p*-values are bounded by this maximum.

| | | AA | AC | AG | AT | CA | CC | CG | CT |
|---|---|---|---|---|---|---|---|---|---|
| Genome | *z* -value | 64.2 | 75.2 | 61.6 | 45.1 | 77.2 | 83.5 | 125.9 | 61.6 |
| | *p* -value | <4.6E-892 | <3.8E-1224 | <1.3E-810 | **<3.4E-442** | <3.6E-1290 | <1.3E-810 | <1.3E-810 | <1.3E-810 |
| Intergenic | *z* -value | 50.3 | 57.5 | 48.1 | 34.8 | 60.2 | 64.7 | 92.6 | 48.1 |
| | *p* -value | <2.2E-545 | <4.5E-708 | <8.2E-503 | **1.9E-264** | <2.5E-784 | <4.6E-892 | <1.0E-1840 | <8.2E-503 |
| Exons | *z* -value | 9.7 | 12.4 | 8.4 | 8.5 | 11.8 | 15.2 | 26.6 | 8.4 |
| | *p* -value | 1.1E-21 | 2.4E-34 | **1.6E-16** | 8.8E-17 | 3.9E-31 | 1.5E-51 | 6.7E-155 | 1.6E-16 |
| Introns | *z* -value | 24.5 | 29.0 | 22.5 | 17.8 | 30.5 | 28.9 | 53.3 | 22.5 |
| | *p* -value | 7.8E-132 | 2.4E-184 | 3.0E-111 | **3.6E-70** | 2.3E-203 | 9.6E-183 | <1.6E-612 | 3.0E-111 |
| | | GA | GC | GG | GT | TA | TC | TG | TT |
| Genome | *z* -value | 77.3 | 72.7 | 83.5 | 75.2 | 92.5 | 77.3 | 77.2 | 64.2 |
| | *p* -value | <3.6E-1290 | <2.3E-1128 | <1.3E-810 | <3.8E-1224 | <1.0E-1840 | <3.6E-1290 | <3.6E-1290 | <4.6E-892 |
| Intergenic | *z* -value | 58.9 | 56.0 | 64.7 | 57.5 | 79.9 | 58.9 | 60.2 | 50.3 |
| | *p* -value | 4.5E-733 | <1.5E-683 | <4.6E-892 | <4.5E-708 | <6.1E-1358 | 4.5E-733 | <2.5E-784 | <2.2E-545 |
| Exons | *z* -value | 17.4 | 12.5 | 15.2 | 12.4 | 12.1 | 17.4 | 11.8 | 9.7 |
| | *p* -value | 5.3E-67 | 6.4E-35 | 1.5E-51 | 2.4E-34 | 9.0E-33 | 5.3E-67 | 3.9E-31 | 1.1E-21 |
| Introns | *z* -value | 28.0 | 30.0 | 28.9 | 29.0 | 39.5 | 28.0 | 30.5 | 24.5 |
| | *p* -value | 3.5E-171 | 2.0E-196 | 9.6E-183 | 2.4E-184 | <1.1E-332 | 3.5E-171 | 2.3E-203 | 7.8E-132 |

**Table S6.** The *z*- and *p*-values from the two-proportion *z*-test on the difference in SNP rates between tissue-differentially- and hypomethylated CpG site blocks, and between hypo- and hypermethylated CpG site blocks. CpG site blocks are categorized into three methylation states: hypomethylated in both of the two tissue types, hypermethylated in both, and tissue-differentially-methylated between the two tissue types.

| | blastulae vs. liver | testes vs. liver |
|---|---|---|
| tissue-differentially- vs. hypomethylated CpG site blocks | | |
| z-value | 25.8 | 19.5 |
| p-value | 3.9E-146 | 2.3E-83 |
| hypo- vs. hypermethylated CpG site blocks | | |
| z-value | 153.9 | 171.6 |
| p-value | <1E-2170 | <1E-2170 |

**Table S7.** Comparisons of identity or difference between methylation states (hypo- or hypermethylated) of CpG blocks among pairs of two tissues. As shown in Fig. 1D, CpG site block methylation states are highly similar in blastulae and testes, whereas the characteristics of CpG site block methylation in liver are somewhat different (Fig. 3A). We therefore investigated the pair of blastulae vs. liver, and the pair of testes vs. liver. **(A)** The *z*- and *p*-values from a two-proportion *z*-test on the difference between dinucleotide substitution rates in regions differentially methylated between tissues and hypomethylated regions in each pair of two tissues. The maximum *p*-value in each category is indicated in bold type. All *p*-values except those for CC and GG are significantly small ($< 5 \times 10^{-4}$). **(B)** The *z*- and *p*-values from a two-proportion *z*-test on the difference between dinucleotide substitution rates in hypo- and hypermethylated regions in each pair of two tissues.

A. The *z*- and *p*-values from a two proportion z-test on the difference between dinucleotide substitution rates in tissue-differentially-methylated regions and hypomethylated regions in each pair of two tissues

| | | AA | AC | AG | AT | CA | CC | CG | CT |
|---|---|---|---|---|---|---|---|---|---|
| medaka-blastula | *z*-value | -18.7 | -18.0 | -19.1 | -8.0 | -23.2 | -2.6 | 24.2 | -19.1 |
| vs liver | *p*-value | 2.0E-77 | 9.4E-72 | 1.9E-80 | 7.0E-15 | 2.7E-118 | **1.3E-02** | 6.9E-128 | 1.9E-80 |
| medaka-testes vs | *z*-value | -17.0 | -13.6 | -16.4 | -4.8 | -17.8 | -0.9 | 24.3 | -16.4 |
| liver | *p*-value | 1.4E-63 | 3.7E-41 | 2.6E-59 | 3.3E-06 | 4.9E-70 | **2.6E-01** | 1.8E-129 | 2.6E-59 |

| | | GA | GC | GG | GT | TA | TC | TG | TT |
|---|---|---|---|---|---|---|---|---|---|
| medaka-blastula | *z*-value | -13.8 | 3.6 | -2.6 | -18.0 | -9.3 | -13.8 | -23.2 | -18.7 |
| vs liver | *p*-value | 1.5E-42 | 6.6E-04 | **1.3E-02** | 9.4E-72 | 4.6E-20 | 1.5E-42 | 2.7E-118 | 2.0E-77 |
| medaka-testes vs | *z*-value | -8.8 | 3.7 | -0.9 | -13.6 | -7.7 | -8.8 | -17.8 | -17.0 |
| liver | *p*-value | 8.8E-18 | 5.0E-04 | **2.6E-01** | 3.7E-41 | 6.4E-14 | 8.8E-18 | 4.9E-70 | 1.4E-63 |

B. The *z*- and *p*-values from a two proportion z-test on the difference between dinucleotide substitution rates in hypo- and hypermethylated regions in each pair of two tissues

| | | AA | AC | AG | AT | CA | CC | CG | CT |
|---|---|---|---|---|---|---|---|---|---|
| medaka-blastula | *z*-value | 74.6 | 85.7 | 70.5 | 49.8 | 87.4 | 89.6 | 131.3 | 70.5 |
| vs liver | *p*-value | <8.6E-1192 | <1.2E-1571 | <1.0E-1066 | **<6.9E-524** | <2.3E-1646 | <8.5E-1723 | <2.6E-2174 | <1.0E-1066 |
| medaka-testes vs | *z*-value | 82.8 | 93.8 | 80.4 | 58.5 | 98.7 | 96.3 | 133.7 | 80.4 |
| liver | *p*-value | <7.7E-1463 | <6.7E-1881 | <1.8E-1392 | **<4.5E-733** | <2.6E-2088 | <4.9E-2004 | <2.6E-2174 | <1.8E-1392 |

| | | GA | GC | GG | GT | TA | TC | TG | TT |
|---|---|---|---|---|---|---|---|---|---|
| medaka-blastula | *z*-value | 85.8 | 77.7 | 89.6 | 85.7 | 89.3 | 85.8 | 87.4 | 74.6 |
| vs liver | *p*-value | <1.2E-1571 | <3.5E-1290 | <8.5E-1723 | <1.2E-1571 | <8.5E-1723 | <1.2E-1571 | <2.3E-1646 | <8.6E-1192 |
| medaka-testes vs | *z*-value | 106.2 | 82.1 | 96.3 | 93.8 | 111.4 | 106.2 | 98.7 | 82.8 |
| liver | *p*-value | <2.6E-2174 | <7.7E-1463 | <4.9E-2004 | <6.7E-1881 | <2.6E-2174 | <2.6E-2174 | <2.6E-2088 | 7.7E-1463 |

**References**

George W. Snedecor WGC. 1989. *Statistical Methods*. Iowa State University Press.

Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* **20**(2): 265-272.

Sasaki S, Mello CC, Shimada A, Nakatani Y, Hashimoto S, Ogawa M, Matsushima K, Gu SG, Kasahara M, Ahsan B et al. 2009. Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* **323**(5912): 401-404.