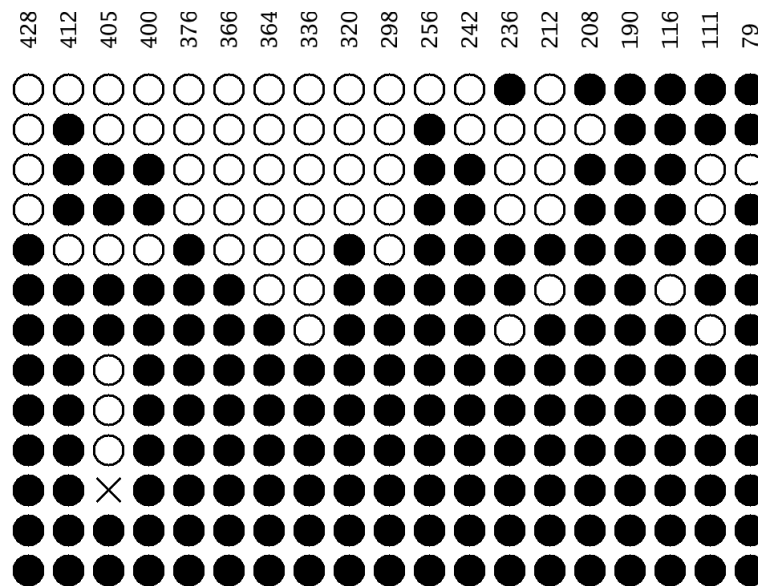


### QUMA tool for bisulfite sequencing analysis

June 25, 2008 – Bisulfite sequencing is a technology for determining the methylation status of stretches of genomic DNA, one of the most widely studied forms of epigenetic marking. The technique relies on the interesting fact that when DNA is treated with the ion  $\text{HSO}_3^-$ , known as bisulfite, it causes all cytosine residues in the sequence to convert into the related nucleobase, uracil; but with one important exception: methylated cytosines remain unchanged. This makes it possible to determine the methylation status of specific sites in the treated sequence with high confidence. One of the main limiting factors in the application of bisulfite sequencing to basic and medical research, however, remains the analysis of the data generated by the DNA sequencer, which requires multiple steps of processing and refinement, including subjective quality checks that may introduce researcher bias.



Sample output automatically generated by QUMA's user-configurable image generator

That all may change with the release of QUMA (quantification tool for methylation analysis), a new software tool for bisulfite sequencing analysis developed by Yuichi Kumaki and colleagues in the Laboratory for Mammalian Epigenetic Studies (Masaki Okano; Team Leader), which addresses some of the shortcomings of previous such applications. The software, which is described in an article in *Nucleic Acids Research*, is available for download under a Creative Commons License from: <http://quma.cdb.riken.jp/top/index.html>.

"I started developing QUMA because I had seen and experienced myself how much time was required for tasks like aligning sequences by hand and generating output for figures," says Kumaki. "I think that people studying DNA methylation will find that this tool will reduce the amount of time they spend on these routine tasks."

QUMA (pronounced koo-ma), boasts a number of features geared toward making bisulfite sequencing studies simpler. The user interface is easy to use and modify, and nearly all of the output data can be downloaded and configured to user specs. It provides a one-stop service for virtually all of the data-processing functions needed to analysis the methylation of sequences, such as sequence alignment, trimming, quality checks, visualization and statistical analysis. This represents a significant improvement over previous software that required, for example, manual alignment and trimming of sequences prior to analysis. QUMA puts quality

RIKEN Center for Developmental Biology (CDB)  
2-2-3 Minatojima minamimachi, Chuo-ku, Kobe 650-0047, Japan

control in the hands of the user with adjustable cutoff parameters, which can be shared with analysis data, increasingly their reproducibility. And, for users who wish to work with sensitive data, the entire software package can be run over a local network from a CD-ROM.

"We're looking forward to receiving feedback from the DNA methylation community on ways to improve the program, and suggestions for new features, such as the ability to analyze methylation status of individual alleles," says Kumaki. "The code is entirely open source, and distributed free via our website, so we hope that this finds a broad user base in the basic and medical research communities."