Advanced DNA methylation analysis for verification of Infinium® HumanMethylation450 BeadChip data

Gerald Schock, Frank Reinecke, Thea Rütjes, Dorothee Unt, Claudia Kappmeier, Norbert Hochstein, Ralf Peist

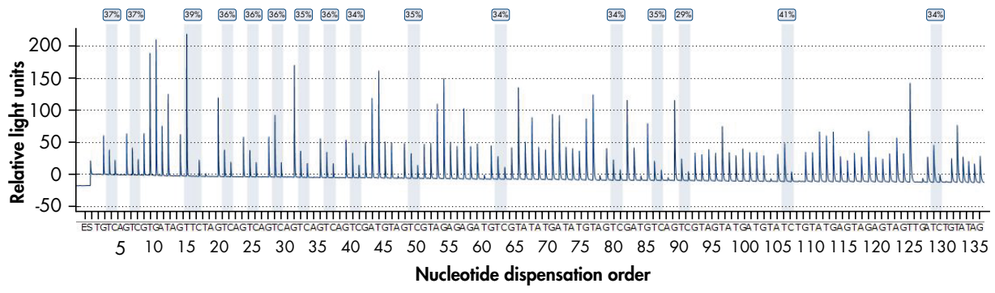
QIAGEN GmbH, QIAGEN Strasse 1, 40724 Hilden, Germany

Introduction

Genome-wide methylation data help identify regions that are important in cancer research. Hybridization arrays allow the screening of a multiple CpG sites for a good overview of potential differentially methylated regions. The widely used Infinium HumanMethylation450 BeadChip assay enables the analysis of over 480,000 different CpG sites in parallel. However, such arrays contain only a subset of all CpG sites. In research on cancer and other diseases featuring heterogeneous CpG methylation, it can be important to analyze additional sites in greater detail and exclude the influence of SNPs, which can interfere with CpG methylation results [1].

A cross-validation study [2] showed very good concordance between a HumanMethylation450 BeadChip array and Pyrosequencing® data, confirming that Pyrosequencing is highly suited for validating array results.

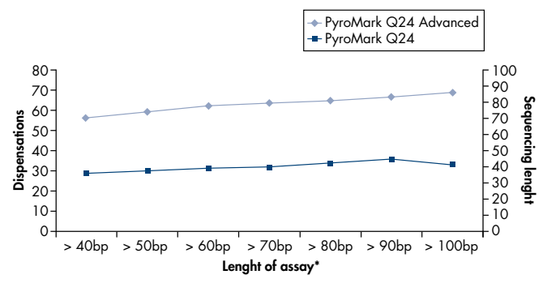
Sequence to analyze:
GYGYGTTTGGGATTGTTTYGGYGYGTYGGGYGYGYGAATTTGTGYGTAAGGGAGAAGAGYGTTTATTATGGAATATGGTTT
GYGGGAGGYGGGYGTAGTTATGATGATYGTAATTGGAGTAGGAGTAAGTTTTTATYGTTTATAGTTATTYG



Analysis of 16 CpG sites in a long sequence run. The PyroMark® Q24 Advanced increases both read length and reliability of methylation analysis at positions later in the sequence. This example demonstrates 135 nucleotide dispensations and the accurate analysis of 16 different CpG positions in a single PyroMark Q24 Advanced CpG reaction. The PyroMark instrument delivers highly reproducible real-time quantification of methylation frequencies at individual CpG sites.

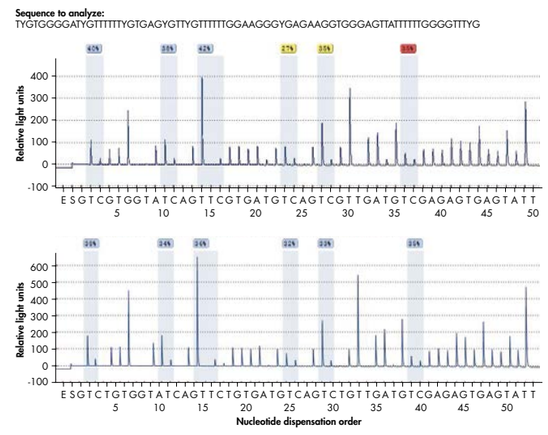
Advanced Pyrosequencing

The new PyroMark Q24 Advanced system overcomes the issue of sequence length limitations that previously hampered Pyrosequencing. It provides sequence reads typically twice as long as those from earlier PyroMark systems. It also reduces background noise in the Pyrosequencing reaction, enabling improved quantification even at sites distant from the sequencing start and after the homopolymer T sequences commonly found in bisulfite-converted DNA.



Comparison of PyroMark Q24 and PyroMark Q24 Advanced performance. 264 Pyrosequencing assays were run in duplicate and classified according to the maximum theoretical sequencing length, as defined by the amplicon length and sequencing primer position. Left axis: Average number of passed dispensations. Right axis: Average sequencing length. PyroMark Q24 Advanced runs typically show passed dispensations twice as long as those from PyroMark Q24 runs.

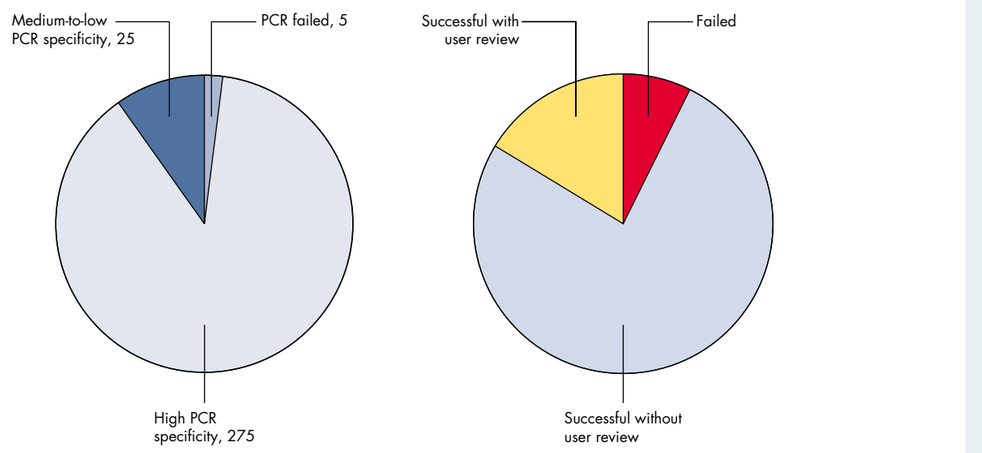
* All assays of the indicated length and above represented by this data point.



Improved quantification of CpG methylation in homopolymers. CpG methylation directly following a T homopolymer or between T homopolymers was quantified by the ratio of converted C nucleotides to unconverted C nucleotides following bisulfite treatment. The PyroMark Q24 Advanced enables accurate quantification of CpG position after or between T homopolymers.

Predesigned CpG assays

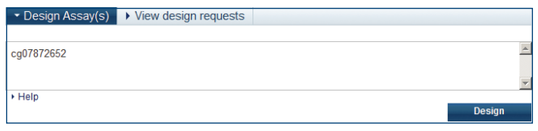
PyroMark CpG Assays enable methylation analysis of specific targets across the genome using carefully tailored algorithms to generate optimal assays. To evaluate the design success of the algorithm, a study was performed using 310 randomly selected genes. PCR success was 98% and Pyrosequencing success was over 92%. The upgraded assay database now includes over 84,000 individual assays for gene-specific human, mouse, and rat CpG sites. Predesigned assays for validating results from methylation arrays such as the HumanMethylation450K BeadChip array have also been developed.



High Pyrosequencing success rate obtained with predesigned PyroMark CpG Assays. The performance of PyroMark CpG Assays was evaluated for 310 randomly selected genes. 305 assays (98%) yielded the predicted PCR amplicon, usable as template for subsequent Pyrosequencing. Of these, 281 yielded successful Pyrosequencing results, with 229 (75.1%) requiring no further user review and 52 (17%) successful after user review. Only 24 (7.9%) of the evaluated assays failed.

PyroMark CpG Assay Design Tool

The new PyroMark CpG Assay Design Tool offers a web interface to facilitate the design of PyroMark CpG Assays for targets of Infinium HumanMethylation450 BeadChip arrays. Detailed information is provided for each assay to help the user to determine the number and position of the analyzed CpG sites. Information about the amplicon length supports the decision on which assay should be used if DNA quality is limited, e.g. for FFPE tissue sample material.



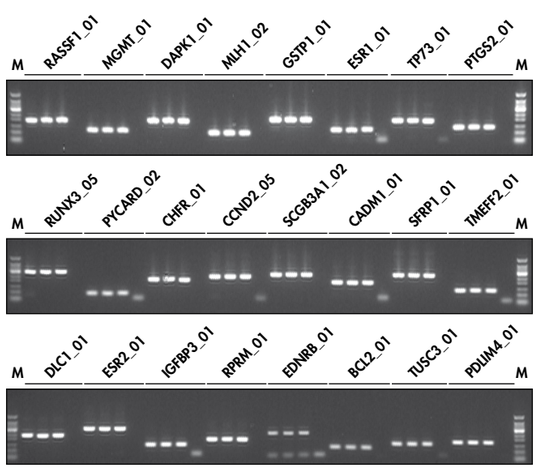
PyroMark CpG Assay design web interface. Assay design for HumanMethylation450 BeadChip array targets starts after entry of the CG#. Assays for multiple targets can be designed in a single request.

Assay design results. A list of available or newly designed assays is displayed after requesting a design. PyroMark platform-specific assay setup files can be downloaded for each assay. The assay setup files make all of the assay-specific information available for the instrument software

Design algorithm and assay optimization

PCR primer specificity is a prerequisite for successful methylation quantification by Pyrosequencing. The design process includes:

1. Template sequence identification for a the target using the unique CpG locus identifier from the Illumina CG database [3].
2. Identification of all common variations using data from dbSNP [4]. These positions will be excluded for primer ends.
3. In-silico bisulfite conversion and exclusion of all variant positions, such as CpG sites.
4. Design of PCR primers according to maximal coverage of the target sequence and optimal melting temperature [5].
5. Virtual amplicon-based selection of forward and reverse primers from all candidates.
6. Verification of the primer pair specificity via a genome-wide search for potential additional priming sites of the selected pair in the bisulfite-treated genomic background sequence.
7. Sequencing primer design according to the nearest position to the target sequence and optimal melting temperature.



Primer design and PCR amplification are optimized for Pyrosequencing. A tailored design algorithm ensures highly specific primer design. PCR was carried out in triplicate using the PyroMark PCR Kit and bisulfite-treated leukocyte DNA as the template. 5 µl of PCR product was loaded onto a 2% agarose gel. **M**: 50 bp marker.

Conclusion

Hybridization arrays such as the Infinium HumanMethylation450 BeadChip from Illumina® enable high-throughput discovery of variations in DNA methylation, but the results must be experimentally verified. Pyrosequencing is a cost-effective technique that is highly suited for this purpose.

The new PyroMark Q24 Advanced system provides improved quantification of CpG methylation and sequence reads that are typically twice as long as those from previous PyroMark systems. Furthermore, the new PyroMark CpG Assay Design web interface offers predesigned PyroMark CpG Assays for validating results from HumanMethylation450 BeadChip experiments without the need to design a separate assay for each individual target. Using a tailored design algorithm, approximately 600,000 assays can be designed for the HumanMethylation450 BeadChip targets. The dedicated algorithm enables a Pyrosequencing success rate of over 92%.

References

1. Price, E.M. et al. (2013) Additional annotation enhances potential for biologically relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip Array. Epigenetics & Chromatin **6 [1]**, 4 doi:10.1186/1756-8935-6-4
2. Roessler, J., Ammerpohl, O., Gutwein, J., Hasemeier, B., Anwar, S.L., Kreipe, H., and Lehmann, U. (2012) Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Research Notes **5**, 210
3. Illumina, Inc. 2010, Technical Note Epigenetics: CpG Loci Identification. Illumina, Inc.
4. Sherry, S.T., et al. (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. **29**, 308
5. Reinecke, F. and DiCarlo, J. (2013) Assay Design for CpG Methylation Analysis using Pyrosequencing (in preparation)

For more information, visit www.qiagen.com.

1080315 01/2014