

Infinium Mouse Methylation GenomeStudio Manifest Column Headings

Apr 28 2021

Below are detailed descriptions of the Infinium Mouse Methylation BeadChip GenomeStudio manifest file column headings.

IlmnID: Unique identifier from the Illumina CG database with four added alphabetic and numeric characters to denote top or bottom strand (T/B), converted or opposite strand (C/O), Infinium probe type (1/2), and the number of synthesis for representation of the probe on the array (1,2,3,...,n).

Name: The locus target identifier (cg, ca, cc, ct, cl, rs) followed by an eight-digit code that relates to the probe sequence. If an eight-digit code has not yet been generated, standard genomic coordinates follow the locus target identifier.

AddressA_ID: For Infinium I bead types, this is the Address ID for the probe specific for the A allele, which is the unmethylated allele. For Infinium II bead types, the Address ID for the probes used for both A and B alleles (in this case, AddressB_ID and AlleleB_ProbeSeq columns are empty).

AlleleA_ProbeSeq: The sequence of the probe identified in the AddressA_ID column.

AddressB_ID: For Infinium I bead types, the address ID for the probe specific for the B allele, which is the methylated allele.

AlleleB_ProbeSeq: For Infinium I bead types, the sequence of the probe identified in the AddressB_ID column.

Next_Base: For Infinium I probes, the nucleotide immediately following the CpG. Blank for Infinium II.

Color_Channel: For Infinium I probes, the color channel of the "Next_Base" signal.

Col: The intended color channel for the probes. The red and green are abbreviated to "R" and "G", respectively.

Probe_Type: Either "ctl", "cg", "ch", "mu", "rp", or "rs" to denote control, CpG, CHG, multi-unique, repetitive element, or SNP probes. Multi-unique probes are probes that map perfectly to multiple sites in the mouse reference genome outside of defined repetitive element regions.

Strand: The Forward (F) or Reverse (R) designation of the Design Strand. *Note: in methylation manifest files, the Forward Strand = the genomic Plus (+) Strand and the Reverse Strand = the genomic Minus (-) Strand. In this context, Forward and Reverse ARE NOT EQUIVALENT to the Forward and Reverse Strand designations originating from reference genome build or as given in Infinium Genotyping manifests.

Strand_TB: The Top (T) or Bottom (B) designation of the Design Strand using Illumina's standard designation for sequences harboring ambiguous SNPs.

Strand_CO: The Converted (C) or Opposite (O) designation of the Design Strand which relates to whether the probe is designed to the strand this is originally converted by bisulfite or the opposite strand that arises after random primer amplification of DNA.

Infinium_Design_Type: Infinium I (2 probes/locus) or Infinium II (1 probe/locus).

Rep_Num: This value depends on how many times a probe was synthesized for representation on the array. For instance, if the probes was synthesized three separate times, the probe generated by the first synthesis would be denoted as "1", while the probe generated from the second synthesis would be denoted as "2", and the probes generated from the third synthesis would be denoted as "3".

CHR: Chromosome containing the CpG (from build mm10).

MAPINFO: Chromosomal coordinates of the CpG (from build mm10).

Species: Name of the species for which a given probes was designed to target.

Genome_Build: Genome Build of the consensus mouse genome.

SourceSeq: The original, genomic sequence used for probe design after bisulfite conversion.

Forward_Sequence: Plus (+) strand sequence (5'-3') flanking the CG.

Top_Sequence: The sequence of the designated TOP strand using Illumina's technique for determining TOP/BOT strands for genotyping, but for the CpG dinucleotide in question rather than for a genetic SNP.

Genome_Build_NCBI: NCBI genome build ID for consensus mouse genome (GRCm38).

N_Shelf: Genomic coordinates of a CpG Island North Shelf, where the array has targeted a CpG within the shelf. The definition of a North Shelf is the region [4000, 2000) base pairs upstream of a CpG Island start site.

N_Shore: Genomic coordinates of a CpG Island North Shore, where the array has targeted a CpG within the shore. The definition of a North Shore is the region [2000, 0) base pairs upstream of a CpG Island start site.

CpG_Island: Genomic coordinates of a CpG Island, where the array has targeted a CpG within the island. CpG Islands are regions greater than 200 base pairs in length with GC content of 50% or greater and have a ratio of >0.6 for the observed number of CG dinucleotides to the expected number considering the total number of G & C bases in the genome segment.

CpG_Island_Chrom: Chromosome where the array has targeted a CpG within a CpG island on the chromosome.

CpG_Island_chromStart: Genome coordinate that denotes the start site of a CpG Island where the array has targeted a CpG within the island.

CpG_Island_chromEnd: Genome coordinate that denotes the stop site of a CpG Island where the array has targeted a CpG within the island.

CpG_Island_Length: Overall nucleotide length of the CpG Island in which the array has targeted a CpG within the island.

CpG_Island_cpgNum: The number of CpGs within a given CpG Island in which the array has targeted a CpG within the island.

CpG_Island_gcNum: The number of G and C bases within a CpG Island in which the array has targeted a CpG within the island.

CpG_Island_perCpg: The percent of CpG dinucleotides as compared to the total bases in a CpG island, where the array has targeted a CpG within the island.

CpG_Island_perGc: The percent of G and C bases within a given CpG island sequence, where the array has targeted a CpG within the island.

CpG_Island_obsExp: The ratio of observed to expected CpG dinucleotides within a CpG island in which the array has targeted a CpG within the island.

S_Shore: Genomic coordinates of a CpG Island South Shore, where the array has targeted a CpG within the shore. The definition of a South Shore is the region [0, 2000) base pairs downstream of the stop site of a CpG Island.

S_Shelf: Genomic coordinates of a CpG Island South Shelf, where the array has targeted a CpG within the shelf. The definition of a North Shelf is the region [2000, 4000) base pairs downstream of the stop site of a CpG Island.

MFG_Change_Flagged: Indicates whether Illumina has detected changes in the functional performance of a specific probe resulting from a change in manufacturing, or if probes generated in different synthesis events function in an unanticipated manner during internal testing. Usually listed as "False" to indicate that no change in functional performance has been detected due to a given manufacturing change. Listed as "True" to indicate if an effect to the functional performance has been detected after a manufacturing change.