For reprint orders, please contact: reprints@futuremedicine.com

Reproducibility-optimized detection of differential DNA methylation

Veronika Suni¹, Fatemeh Seyednasrollah¹, Bishwa Ghimire¹, Sini Junttila¹, Asta Laiho¹ & Laura L Elo^{*,1}

¹Turku Bioscience Centre, University of Turku & Åbo Akademi University, FI-20520 Turku, Finland *Author for correspondence: laura.elo@utu.fi

Aim: DNA methylation is a key epigenetic mechanism regulating gene expression. Identifying differentially methylated regions is integral to DNA methylation analysis and there is a need for robust tools reliably detecting regions with significant differences in their methylation status. **Materials & methods:** We present here a reproducibility-optimized test statistic (ROTS) for detection of differential DNA methylation from high-throughput sequencing or array-based data. **Results:** Using both simulated and real data, we demonstrate the ability of ROTS to identify differential methylation between sample groups. **Conclusion:** Compared with state-of-the-art methods, ROTS shows competitive sensitivity and specificity in detecting consistently differentially methylated regions.

First draft submitted: 1 October 2019; Accepted for publication: 25 February 2020; Published online: 4 June 2020

Keywords: differential methylation • DNA methylation • reduced representation bisulfite sequencing • reproducibility • ROTS

DNA methylation is a major epigenetic mechanism that plays a central role in gene regulation. It affects a wide spectrum of biological states from normal development and aging to complex diseases such as cancer [1,2] and diabetes [3,4]. Among the available platforms for DNA methylation detection, methods based on bisulfite conversion are the most popular, including both microarray- and high-throughput sequencing-based techniques [5]. For methylation microarray data, Illumina Infinium platform has been widely used, including for instance more than 10,000 samples in the Cancer Genome Atlas (TCGA) and around 40,000 human samples in the Gene Expression Omnibus [6]. With high-throughput sequencing techniques, whole-genome bisulfite sequencing provides whole-genome coverage with single-base resolution, whereas reduced representation bisulfite sequencing (RRBS) is a cost-effective alternative, covering genomic areas enriched with a high CpG content (~85% of the CpG islands in the human genome) [7,8].

The primary goal of DNA methylation analysis is to identify systematic differences between groups of samples, such as cases versus controls. To enable reliable detection of differentially methylated regions (DMRs) for further validation and interpretation, robust tools are needed to rank DMRs on the basis of their significance. Although computational tools have already been developed for differential methylation analysis [9–13], the field lacks consensus on the best tools. While a recent comparison highlighted the differences between the methods, none of the compared methods could be considered as a generally applicable leading methodology [14].

To address the need for a method enabling robust identification of DMRs, we demonstrate here the utility of our reproducibility-optimized test statistic (ROTS) [15] for differential DNA methylation analysis. ROTS was originally designed for differential gene expression analysis and its good performance has been shown in a number of gene and protein expression studies [15–18]. Here, we adapted ROTS to DNA methylation studies and showed its benefits in both sequencing- and array-based DNA methylation data.

Materials & methods

Sequencing datasets & preprocessing

Simulated DNA methylation sequencing data were generated as described in [14] using dataSim2.R downloaded from the online repository containing the source code and additional resources for the article. The read coverage



of the simulated data was modeled by a binomial distribution and the methylation background followed beta distribution [14]. A total of 12 different simulation scenarios were considered, including methylation change of 10, 15 or 25%, number of sites of 5000 or 50,000 and proportion of true positives of 5 or 20%. Five samples were simulated in the control group and five samples in the case group.

The real DNA methylation sequencing data were from a previous RRBS study on human pluripotent stem cells [19]. The data included six samples before and six samples after spontaneous transformation to abnormal karyotypes, referred to as normal and abnormal samples, respectively. The DNA libraries were sequenced using 50 bp reads on the Illumina HiSeq2000 or HiSeq2500 platform. The data were downloaded from www.ncbi.nlm.nih .gov/bioproject under the accession number PRJNA310646.

The quality control and preprocessing of the sequenced reads were performed using Trim Galore! 0.4.1 along with Cutadapt 1.9.dev. The nondefault Trim Galore! parameters used were –stringency = 3 (to remove adapter parts overlapping with three or more bases) and –rrbs (to remove filled-in cytosines). The preprocessed reads were mapped to the hg19 reference genome using Bismark version 0.10 [20] with Bowtie 2 version 2.0.5 [21] using default parameters. The methylation calling was performed using the Bismark methylation extractor module in default mode, resulting in base-level methylation values representing the proportion of methylation in range 0–100. Only CpG sites with a methylation coverage of at least ten reads were used in the downstream analysis. Base-pair resolution cytosine methylation levels were further summarized to 100 bp windows by taking the median across the base-level methylation level values within each window.

For statistical testing, we required that a feature (single site or window) had at least three nonmissing values in both sample groups. Additionally, features with a constant value across all samples when binarized (values below 50 converted to zero and values above or equal to 50–100) were excluded, leaving 33,282 single sites or 18,238 windows for statistical testing.

Array-based dataset & preprocessing

The array-based DNA methylation data were downloaded from TCGA database via the National Cancer Institute (NCI) Genomic Data Commons data portal under the project named TCGA Kidney Renal Clear Cell Carcinoma (TCGA-KIRC). The data included baseline tumor samples from 398 patients diagnosed with clear cell renal cell carcinoma (ccRCC) together with clinical information, involving 132 metastatic and 266 nonmetastatic patients. Two Illumina Infinium DNA methylation platforms, HumanMethylation27 (HM27) BeadChip and HumanMethylation450 (HM450) BeadChip, were used to obtain the methylation profiles [6].

The microarray data were preprocessed by TCGA using the R package *methylumi* version 2.22.0 [22]. Briefly, the DNA methylation level for each CpG locus was summarized as a beta value ranging from zero to one, representing the proportion of methylation. Probes corresponding to loci having nonsignificant quality values (p > 0.05) were filtered out, leaving 21,073 probes for statistical testing.

ROTS

The ROTS is a data adaptive approach. Instead of making any *a priori* assumptions about the characteristics of the data, ROTS utilizes a family of t-type statistics (d_{α}) to optimize the reproducibility $R_k(d_{\alpha})$ of the *k* top-ranked features in pairs of bootstrapped datasets [16]. More specifically, ROTS maximizes the reproducibility *z*-statistic $Z_k(d_{\alpha})$:

$$Z_k(d_\alpha) = \frac{R_k(d_\alpha) - R_k^0(d_\alpha)}{s_k(d_\alpha)}$$

over parameters α and the top list size k. Here, $R_k(d_\alpha)$ and $R_k^0(d_\alpha)$ denote the reproducibility of the bootstrapped and randomized data, respectively, and $s_k(d_\alpha)$ is the estimated standard deviation of the bootstrap distribution. The reproducibility $R_k(d_\alpha)$ is defined as the average overlap of the k top-ranked features over B pairs of bootstrap datasets. For each pair b of bootstrap data matrices $(D_1^{(b)}, D_2^{(b)})$ the reproducibility is calculated as

$$R_{k}^{(b)}(d_{\alpha}) = \frac{\#\left\{g | r_{g}(\alpha, D_{1}^{(b)}) \leq k, r_{g}\left(\alpha, D_{2}^{(b)}\right) \leq k\right\}}{k}$$

where $r_g(\alpha, D_i^{(b)})$ denotes the rank of feature g in data $D_i^{(b)}$ with the statistic d_α and #S is the cardinality of set S.

The test statistics d_{α} for a genomic feature of g (here DNA methylation region) is defined as

$$d_{\alpha}(g) = \frac{|\overline{x}_{g}^{i} - \overline{x}_{g}^{j}|}{\alpha_{1} + \alpha_{2}s_{g}},$$

where \overline{x}_{g}^{i} and \overline{x}_{g}^{j} are the average methylation levels of feature g in the experimental conditions i and j and s_{g} represents the estimated standard error and $\alpha_{1} \ge 0$ and $\alpha_{2} \in \{0,1\}$ are the parameters to be optimized.

As input data, ROTS requires a matrix of preprocessed and normalized data with columns representing the samples and rows representing the methylation regions. The ROTS package and a detailed manual are freely available through Bioconductor at http://bioconductor.org/packages/ROTS.

Differential methylation analysis of the sequencing data

Differential methylation analysis of the DNA methylation sequencing data was performed using ROTS, RnBeads and MethylKit using the same input data.

For ROTS (version 1.5.4 on R 3.3.3), the methylation values in range 0–100 were scaled to range zero-to-one, the number of bootstraps was set to B = 1000 and the maximum number of top-ranked features for reproducibility optimization was set to K = 10,000. ROTS reports a top list size k, which gives the highest reproducibility Z-score. If the reported k is close to the value of the parameter K, the maximum top list size might have been too small and increasing it should be considered. It is recommended that the value of K should be considerably higher than the number of features expected to be significantly different between the sample groups [15]. As the reported k was 480 and 525 for the human embryonic stem cell RRBS data and array-based methylation data, respectively, the selected value of K was considerably larger than the reported k, as recommended.

For RnBeads (version 1.6.1 on R 3.3.3), the parameter *exploratory.region.profiles* was used to consider the prefiltered 100 bp windows as input. The following additional parameters were used: *filtering.high.coverage.outliers* = TRUE, *filtering.sex.chromosomes.removal* = FALSE, *filtering.coverage.threshold* = 10, *filtering.snp* = 'no', *differential.site.test.method* = 'limma'. RnBeads uses hierarchical linear models implemented in R package Limma to calculate p-values, which are subjected to multiple-testing correction using the false discovery rate (FDR) method.

For MethylKit (version 1.1.7), its built-in function of tiling window analysis was used for the initial window-wise analysis with *win.size* = 100 and *step.size* = 100. Bases having read coverage below ten and bases with more than 99.9th percentile of coverage in each sample were removed as a default procedure during the MethylKit run. The window-wise results were then further filtered to cover the same regions that were used as input for the other methods. MethylKit uses logistic regression to calculate p-values and sliding linear model method [23] to adjust the p-values to q-values.

Regions with an FDR below 0.05 and a methylation change larger than 15% between the groups were considered as significantly differentially methylated regions (DMRs). DMRs were ranked based on the p-values calculated using the respective statistical test implemented in each of the tested tools.

Differential methylation analysis of the array-based data

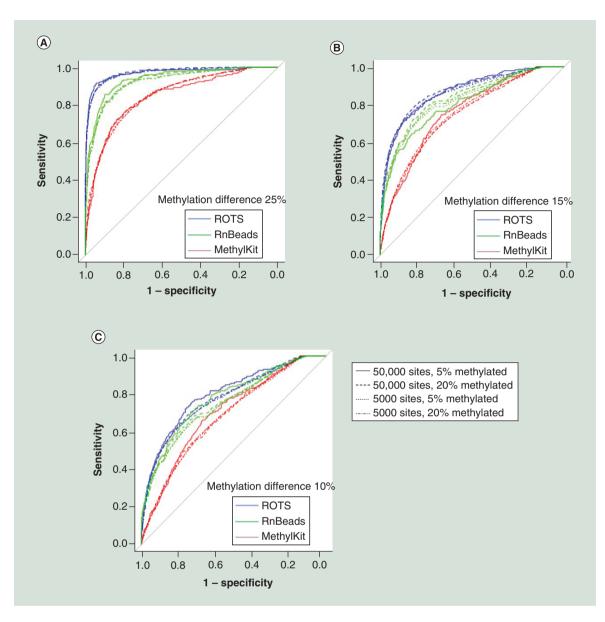
Differential methylation analysis of the array-based DNA methylation data was performed using ROTS (version 1.5.4 on R 3.3.3) and Limma (version 3.32.5). The beta values were further transformed into so called M-values ($\log_2(\frac{beta}{1-beta})$), as suggested previously [24]. For ROTS, the number of bootstraps was set to B = 1000 and the maximum number of top-ranked features for reproducibility optimization to K = 10,000. The differentially methylated probes were ranked based on the p-values calculated using the respective statistical test implemented in the tested tools.

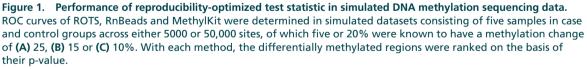
Results

Simulated data

We first evaluated the performance of ROTS on simulated DNA methylation sequencing data, produced similarly as in a previous simulation study [14] and compared it with two popular tools, MethylKit [9] and RnBeads [10]. The simulated datasets consisted of five samples in case and control groups across either 5000 or 50,000 sites, of which five or 20% were known to have a methylation change of 10, 15 or 25%.

As expected, the larger the methylation changes, the better all the methods performed. The receiver operating characteristic curves showed that ROTS confidently assigned DMRs with higher sensitivity and specificity than





ROC: Receiver operating characteristic; ROTS: Reproducibility-optimized test statistic.

both RnBeads and MethylKit; in 11 out of 12 simulation scenarios, the differences were statistically significant (Figure 1; DeLong's test p < 0.05, Supplementary Table 1).

Human embryonic stem cell RRBS data

Next, we used ROTS to reanalyze the previously published RRBS dataset on human embryonic stem cells, consisting of six samples in the normal group and six samples in the abnormal group [19]. As the majority of DMRs reported are in the range of several hundred to a few thousand bases [25], we focused here on 100 bp regions, but differential methylation of single CpG cytosines was also analyzed (Supplementary Figure 1).

ROTS identified 291 DMRs between the normal and abnormal groups with a FDR <0.05 and an absolute methylation change >15%. With MethylKit and RnBeads, the corresponding numbers were 6470 and 133,

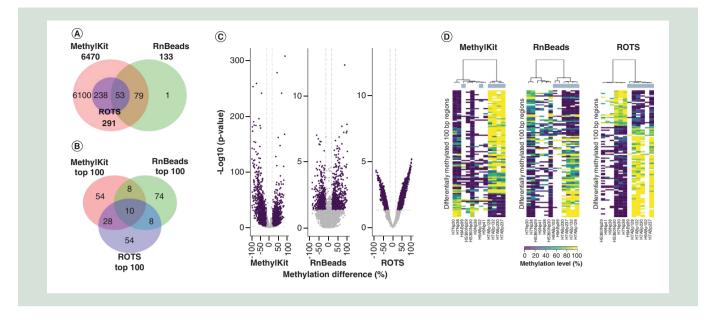


Figure 2. Performance of reproducibility-optimized test statistic in DNA methylation sequencing data on human embryonic stem cells. (**A**) Overlaps of all statistically significantly differentially methylated 100 bp regions (DMRs) detected by MethylKit, RnBeads and ROTS at false discovery rate of 0.05. (**B**) Overlaps of 100 top-ranked DMRs detected by MethylKit, RnBeads and ROTS. (**C**) Volcano plots showing the relationship between the estimated methylation change (x-axis) and its significance (y-axis). The purple dots represent DMRs with a p-value <0.05 and a methylation change >15%. (**D**) Heatmaps showing the 100 top-ranked DMRs identified by MethylKit, RnBeads or ROTS. The colors indicate the methylation level (methylation percentage) of each DMR. Missing values are shown in white. The normal and abnormal samples are marked in blue and white, respectively. (**E–G**) Representative examples of highly-ranked DMRs detected by (**E**) ROTS and MethylKit but not by RnBeads, (**F**) by RnBeads but not by ROTS and (**G**) by MethylKit but not by ROTS or RnBeads. The six abnormal (case, six upper tracks) and six normal (control, six lower tracks) samples are shown as separate tracks. The numbers of bisulfite-converted (blue) and nonconverted (red) cytosines at CpG sites with at least ten reads coverage are shown. Similar plots for all the 100 top-ranked DMRs with each method are shown in Supplementary files 2–4. ROTS: Reproducibility-optimized test statistic.

respectively. All DMRs detected by ROTS were also detected by MethylKit, while only 18% of the ROTSidentified DMRs were detected by RnBeads (Figure 2A). When looking only at the 100 top-ranked DMRs for each method, the overlap of ROTS with MethylKit decreased to 38%, while the overlap of ROTS with RnBeads remained at 18% (Figure 2B). All but one of the DMRs detected by RnBeads were also detected by MethylKit, with the overlap decreasing to 18% when focusing on the 100 top-ranked DMRs. Analysis of single cytosine sites showed similar performance with 365, 10,470 and 317 cytosines detected as differentially methylated by ROTS, MethylKit and RnBeads, respectively (Supplementary Figure 1).

Investigation of the methylation change versus its significance suggested that the ideal 'V'-shape of the volcano plot was more prominent with ROTS than with the other tools (Figure 2C). Especially with RnBeads, the relationship between the methylation change and its significance was considerably distorted. In line with this, a heatmap visualization of the 100 top-ranked DMRs with each method revealed that ROTS had the clearest separation between the sample groups (Figure 2D). Notably, eight out of the nine differentially methylated CpG sites validated in the original study on the basis of their MethylKit analysis [19] were identified as significant at FDR <0.1 with both ROTS and MethylKit, but only one with RnBeads (Supplementary Table 2). The one site that was not detected by ROTS and MethylKit was filtered out due to its low read coverage.

To further explore the quality of the DMRs reported by the tools, we inspected the 100 top-ranked DMRs with each method in detail (Figure 2E–G, Supplementary Figures 2–5, Supplementary files 2–4). The median rank of the 100 top-ranked DMRs with ROTS was 343 and 154 with RnBeads and MethylKit, respectively, while the median rank of the 100 top-ranked DMRs with MethylKit was 601 and 173 with RnBeads and ROTS, respectively and with RnBeads 737 and 475 with MethylKit and ROTS, respectively.

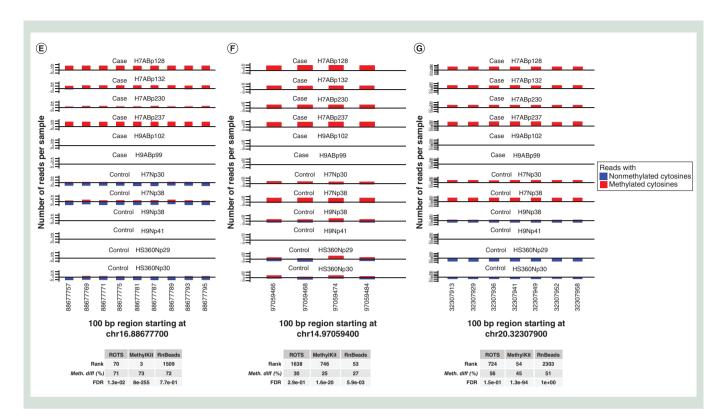


Figure 2. Performance of reproducibility-optimized test statistic in DNA methylation sequencing data on human embryonic stem cells (cont.). (A) Overlaps of all statistically significantly differentially methylated 100 bp regions (DMRs) detected by MethylKit, RnBeads and ROTS at false discovery rate of 0.05. (B) Overlaps of 100 top-ranked DMRs detected by MethylKit, RnBeads and ROTS. (C) Volcano plots showing the relationship between the estimated methylation change (x-axis) and its significance (y-axis). The purple dots represent DMRs with a p-value <0.05 and a methylation change >15%. (D) Heatmaps showing the 100 top-ranked DMRs identified by MethylKit, RnBeads or ROTS. The colors indicate the methylation level (methylation percentage) of each DMR. Missing values are shown in white. The normal and abnormal samples are marked in blue and white, respectively. (E–G) Representative examples of highly-ranked DMRs detected by (E) ROTS and MethylKit but not by RnBeads, (F) by RnBeads but not by ROTS and (G) by MethylKit but not by ROTS or RnBeads. The six abnormal (case, six upper tracks) and six normal (control, six lower tracks) samples are shown as separate tracks. The numbers of bisulfite-converted (blue) and nonconverted (red) cytosines at CpG sites with at least ten reads coverage are shown. Similar plots for all the 100 top-ranked DMRs with each method are shown in Supplementary files 2–4. ROTS: Reproducibility-optimized test statistic.

Array-based DNA methylation data on renal cell carcinoma

After confirming the competitive performance of ROTS with the simulated DNA methylation sequencing data, we assessed its utility in array-based DNA methylation data on ccRCC. Despite the emergence of the sequencing techniques, array-based DNA methylation data still contribute a valuable source of information in the field. Here, we used DNA methylation profiles of 398 ccRCC patients provided by the TCGA to detect DMRs between metastatic and nonmetastatic patients [26]. To benchmark the performance of ROTS, the analysis results were compared with those of Limma [27], which is part of the currently widely used DNA methylation analysis workflows [24,28].

Both ROTS and Limma detected a large number of DMRs at FDR <0.05 (7581 and 9431, respectively), but overall, the methylation changes between the metastatic and nonmetastatic patients were small. Only 296 DMRs had a methylation change of at least 10% and only 24 DMRs had at least 15%; both ROTS and Limma detected all these DMRs as significant at FDR <0.05. When focusing only on 100 top-ranking DMRs with ROTS and Limma, hierarchical clustering grouped the patients into two main clusters, one cluster comprising mainly of metastatic patients and the other mainly of nonmetastatic patients (Supplementary Figure 6).

Runtime

The human embryonic stem cell RRBS dataset was used to study the differences in the runtime of the three methods. The differential methylation analyses were performed with a standard Macintosh laptop computer with a

2.9 GHz Intel Core i7 processor and 16 GB of RAM. MethylKit was the fastest method to run with the differential methylation analysis taking 3 min and 17 s. RnBeads took 5 min and 29 s to run while ROTS' runtime was 9 min 58 s. Although ROTS was the slowest method to run, it should be noted that the preprocessing of sequencing data (including alignment to the reference genome and extraction of the methylation values) is a much more time-consuming and computationally demanding step in the overall analysis than the differential methylation analysis where the differences in runtime between the methods are relatively small.

Discussion

Here, we have studied the suitability of ROTS to differential methylation analysis using simulated and real data and have compared its performance with that of widely-used differential methylation analysis tools. Our results show that ROTS works well with different types of methylation data and they support the conclusion that ROTS is able to robustly identify DMRs between conditions.

In the differential methylation analysis of the RRBS data, MethylKit identified a considerably higher number of DMRs than either ROTS or RnBeads (Figure 2A). Similar behavior has been detected in a previous study, where nine methods for differential methylation analysis were compared with MethylKit identifying over 10,000 DMRs whereas most of the other methods identified 5000 or less DMRs [29]. In concordance with these results, MethylKit DMR detection has been found to have high sensitivity but low specificity [14].

The top 100 DMRs detected by ROTS distinctly separated the normal and abnormal samples, while the results suggest that MethylKit and RnBeads allowed more variation between the replicates compared with ROTS (Figure 2D). Similarly, a detailed inspection of the 100 top-ranked DMRs in each method supported that the DMRs detected by ROTS had high agreement between the biological replicates (Figure 2E, Supplementary file 2). These DMRs were also relatively highly ranked according to RnBeads and MethylKit (median rank 343 and 154, respectively). The top DMRs identified only by MethylKit or RnBeads, instead, tended to show considerable variation between the biological replicates (Figure 2F–G, Supplementary files 3–4).

In the array-based data, both ROTS and Limma reported a large number of DMRs and the differences in methylation between the nonmetastatic and metastatic patients were small. Notably, in the hierarchical clustering of the 100 top-ranking DMRs, ROTS grouped only 23% of the metastatic patients in the nonmetastatic cluster, compared with 34% with Limma. Overall, ROTS suggested a somewhat larger number of patients into the metastatic cluster than Limma (175 and 133 patients with ROTS and Limma, respectively). It should be noted that the original study had some limitations in determining the status of clinical metastasis at the time of the nephrectomy [26] and, hence, some of the patients defined as nonmetastatic may actually have been metastatic or later developed metastatic disease. This can potentially explain also the relatively small methylation changes observed between the groups.

ROTS was originally designed for differential expression analysis of gene expression microarray data and it has previously been shown in an independent study to work well with RNA-sequencing data [30]. The present results show that it performs well also with sequencing- and array-based methylation data. Similarly, previous comparison studies have concluded that methods designed originally for microarray data, such as Limma, are suitable also for sequencing data and often work at least as well as methods developed specifically for sequencing data [31,32]. Among the methods compared in this study, the differential methylation analysis in RnBeads uses Limma, while MethylKit is developed specifically for differential methylation analysis of sequencing data. Limma and ROTS are based on moderated t-statistics whereas MethylKit does not have any distributional assumptions about the normality of the data, but regardless, performs worse. Notably, ROTS p-values are not directly based on any parametric assumptions about the distribution of the data, but they are estimated by permutation.

Conclusion

Our results showed the benefits and versatility of ROTS for effective detection of DNA methylation changes between sample groups. The method is applicable to both high-throughput sequencing and array-based DNA methylation data. Thanks to its data-adaptive approach, ROTS was able to detect accurately group-specific DMRs with high consistency across replicates, as compared with state-of-the-art methods. This is crucial to ensure reproducible findings.

Summary points

- DNA methylation is a key epigenetic mechanism that has been shown to play a role in common diseases like cancer and diabetes.
- Accurate detection of differentially methylated regions is still a major challenge in DNA methylation analysis.
- There is a need for a robust tool that reliably detects significantly differentially methylated regions between sample groups.
- Reproducibility-optimized test statistic (ROTS) is a data-adaptive approach that has previously been shown to yield reliable results in several gene and protein expression studies.
- We demonstrated the applicability of ROTS to differential methylation analysis using both simulated and real data sets.
- ROTS showed good performance with sequencing- as well as array-based methylation data.
- Compared to popular methylation analysis methods ROTS accurately identified differentially methylated regions with high consistency across replicates.
- We believe that ROTS will provide researchers an accurate, reliable and easy-to-use tool for differential methylation analysis.
- ROTS is implemented as an R package and freely available through Bioconductor (www.bioconductor.org).

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/sup pl/10.2217/epi-2019-0289

Acknowledgments

The authors thank the Elo lab for fruitful discussions and comments on the manuscript.

Financial & competing interests disclosure

This work was supported by the European Research Council ERC (grant number 677943), European Union's Horizon 2020 research and innovation program (grant number 675395), Academy of Finland (grant numbers 296801, 304995, 310561, 313343 and 329278), Juvenile Diabetes Research Foundation JDRF (grant number 2-2013-32), Business Finland (grant number 1877/31/2016) and Sigrid Juselius Foundation. Turku Graduate School (UTUGS) also supported this work. Our research is also supported by University of Turku, Åbo Akademi University, Biocenter Finland and ELIXIR Finland node. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Data sharing statement

ROTS is freely available through Bioconductor (https://www.bioconductor.org). The human embryonic stem cell RRBS data were downloaded from BioProject under the accession number PRJNA310646 and the array-based DNA methylation data from TCGA database under the project name TCGA-KIRC.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

References

Papers of special note have been highlighted as: • of interest

- 1. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA methylation in cancer and aging. Cancer Res. 76(12), 3446–3450 (2016).
- 2. Joo JE, Dowty JG, Milne RL *et al.* Heritable DNA methylation marks associated with susceptibility to breast cancer. *Nat. Commun.* 9(1), 867 (2018).
- 3. Paul DS, Teschendorff AE, Dang MA *et al.* Increased DNA methylation variability in type 1 diabetes across three immune effector cell types. *Nat. Commun.* 7, 13555 (2016).
- 4. Rönn T, Ling C. DNA methylation as a diagnostic and therapeutic target in the battle against Type 2 diabetes. *Epigenomics* 7(3), 451–460 (2015).
- 5. Kurdyukov S, Bullock M. DNA methylation analysis: choosing the right method. Biology (Basel) 5(1), 3 (2016).

- 6. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* 45(4), e22 (2017).
- Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6(4), 468–481 (2011).
- 8. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* 33(18), 5868–5877 (2005).
- 9. Akalin A, Kormaksson M, Li S *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 13(10), R87 (2012).
- Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat. Methods 11(11), 1138–1140 (2014).
- 11. Stockwell PA, Chatterjee A, Rodger EJ, Morison IM. DMAP: differential methylation analysis package for RRBS and WGBS data. *Bioinformatics* 30(13), 1814–1822 (2014).
- 12. Park Y, Wu H. Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* 32(10), 1446–1453 (2016).
- 13. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 13(10), R83 (2012).
- 14. Wreczycka K, Gosdschan A, Yusuf D, Grüning B, Assenov Y, Akalin A. Strategies for analyzing bisulfite sequencing data. *J. Biotechnol.* 261, 105–115 (2017).
- A comprehensive review of differential methylation analysis methods and tools.
- 15. Suomi T, Seyednasrollah F, Jaakkola MK, Faux T, Elo LL. ROTS: an R package for reproducibility-optimized statistical testing. *PLoS Comput. Biol.* 13(5), e1005562 (2017).
- Elo LL, Filén S, Lahesmaa R, Aittokallio T. Reproducibility-optimized test statistic for ranking genes in microarray studies. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 5(3), 423–431 (2008).
- Describes the statistical principle behind reproducibility-optimized test statistic.
- Jaakkola MK, Seyednasrollah F, Mehmood A, Elo LL. Comparison of methods to detect differentially expressed genes between single-cell populations. *Brief Bioinform.* 18(5), 735–743 (2017).
- Seyednasrollah F, Rantanen K, Jaakkola P, Elo LL. ROTS: reproducible RNA-seq biomarker detector-prognostic markers for clear cell renal cell cancer. *Nucleic Acids Res.* 44(1), e1 (2016).
- 19. Konki M, Pasumarthy K, Malonzo M et al. Epigenetic silencing of the key antioxidant enzyme catalase in karyotypically abnormal human pluripotent stem cells. Sci. Rep. 6, 22190 (2016).
- 20. Krueger F andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* 27(11), 1571–1572 (2011).
- 21. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10(3), R25 (2009).
- 22. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* 41(7), e90 (2013).
- 23. Wang HQ, Tuominen LK, Tsai CJ. SLIM: a sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. *Bioinformatics* 27(2), 225–231 (2011).
- 24. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. *F1000Res.* 5, 1281 (2016).
- 25. Bock C. Analysing and interpreting DNA methylation data. Nat. Rev. Genet. 13(10), 705-719 (2012).
- 26. Network CGAR. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 499(7456), 43-49 (2013).
- 27. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, Article3 (2004).
- 28. Wilhelm-Benartzi CS, Koestler DC, Karagas MR *et al.* Review of processing and analysis methods for DNA methylation array data. *Br. J. Cancer* 109(6), 1394–1402 (2013).
- 29. Klein HU, Hebestreit K. An evaluation of methods to test predefined genomic regions for differential methylation in bisulfite sequencing data. *Brief Bioinform.* 17(5), 796–807 (2016).
- 30. Soneson C, Robinson M. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods.* 15(4), 255–261 (2018).
- Seyednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. Brief Bioinform. 16(1), 59–70 (2015).
- 32. Rapaport F, Khanin R, Liang Y *et al.* Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol.* 14, 3158 (2013).

Table S1. Performance of ROTS on simulated DNA methylation sequencing data. Areas under the receiver operating characteristic curves (AUC) of ROTS, RnBeads and MethylKit were determined in simulated datasets consisting of five samples in case and control groups across either 5,000 or 50,000 sites, of which 5% or 20% were known to have a methylation change of 25%, 15% or 10%. The differentially methylated regions were ranked on the basis of their p-value. The highest AUC value in each data is highlighted.

Methylation	Total number of	% and number of		MethylKit	RnBeads ROTS		ROTS	DeLong's test, p-value		
change	simulated sites	methylated sites	AUC	95% Confidence interval	AUC	95% Confidence interval	AUC	95% Confidence interval	ROTS vs MethylKit	ROTS vs RnBeads
25%	5,000	5%, 250	0.841	0.814-0.867	0.938	0.923-0.953	0.97	0.957-0.983	2.2 x 10 ⁻¹⁶	1.63 x 10 ⁻³
		20%, 1,000	0.851	0.838-0.865	0.927	0.918-0.936	0.974	0.969-0.978	2.2 x 10 ⁻¹⁶	2.2 x 10 ⁻¹⁶
	50,000	5%, 2,500	0.848	0.84-0.856	0.924	0.918-0.929	0.969	0.966-0.972	2.2 x 10 ⁻¹⁶	2.2 x 10 ⁻¹⁶
		20%, 10,000	0.847	0.843-0.852	0.923	0.92-0.927	0.969	0.967-0.971	2.2 x 10 ⁻¹⁶	2.2 x 10 ⁻¹⁶
15%	5,000	5%, 250	0.756	0.727-0.785	0.805	0.775-0.834	0.867	0.859-0.885	5.83 x 10 ⁻¹⁴	1.27 x 10 ⁻³
		20%, 1,000	0.741	0.724-0.758	0.837	0.823-0.852	0.872	0.843-0.891	2.2 x 10 ⁻¹⁶	4.79 x 10⁻⁴
	50,000	5%, 2,500	0.742	0.733-0.752	0.818	0.809-0.828	0.86	0.852-0.868	2.2 x 10 ⁻¹⁶	5.47 x 10 ⁻¹¹
		20%, 10,000	0.751	0.746-0.756	0.827	0.822-0.831	0.866	0.862-0.87	2.2 x 10 ⁻¹⁶	2.2 x 10 ⁻¹⁶
10%	5,000	5%, 250	0.7	0.668-0.732	0.778	0.747-0.809	0.798	0.769-0.828	1.17 x 10 ⁻⁷	0.344
		20%, 1,000	0.683	0.665-0.701	0.752	0.735-0.77	0.779	0.763-0.796	2.2 x 10 ⁻¹⁶	2.8 x 10 ⁻²
	50,000	5%, 2,500	0.68	0.67-0.69	0.747	0.736-0.757	0.773	0.762-0.783	2.2 x 10 ⁻¹⁶	4.63 x 10⁻⁴
		20%, 10,000	0.683	0.677-0.688	0.748	0.743-0.754	0.774	0.769-0.779	2.2 x 10 ⁻¹⁶	6.56 x 10 ⁻¹¹

Table S2. Performance of ROTS on DNA methylation sequencing data of human embryonic stem cells in terms of the nine differentially methylated CpG sites validated in the original study [19]. The differentially methylated sites and 100 bp regions were ranked on the basis of their p-value. The third site was filtered out due to its low read coverage across the samples.

	Diff. methylated 100 bp regions	MethylKit		RnBeads		ROTS	
Diff. methylated sites		Rank	FDR	Rank	FDR	Rank	FDR
	1. chr1_243651100_243651199	88	2.99 x 10 ⁻⁷⁵	160	0.064	89	0.017
1-2. chr1_243651142,		244	1.23 x 10 ⁻²¹	234	0.014	282	0.038
chr1_243651149		136	7.82 x 10 ⁻²⁶	631	0.145	284	0.038
	2. chr17_40824400_40824499						
3. chr17_40824405		-	-	-	-	-	-
	3. chr11_34460500_34460599	109	1.55 x 10 ⁻⁶⁸	695	0.419	43	0.013
4. chr11_34460572		416	3.53 x 10 ⁻¹⁷	1326	0.410	119	0.017
	4. chr5_178487200_178487299	25	1.49 x 10 ⁻¹²⁹	557	0.342	160	0.024
5-9. chr5_178487247,		306	1.64 x 10 ⁻¹⁹	708	0.182	227	0.031
chr5_178487249,		394	1.38 x 10 ⁻¹⁷	2415	0.767	521	0.069
chr5_178487273,		443	7.10 x 10 ⁻¹⁷	2439	0.774	755	0.099
chr5_178487279,		374	6.56 x 10 ⁻¹⁸	1196	0.362	647	0.088
chr5_178487297		375	6.56 x 10 ⁻¹⁸	1197	0.362	648	0.088

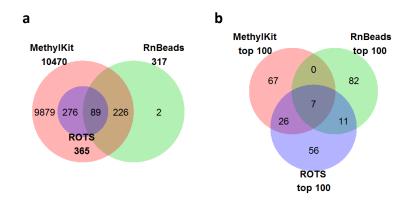


Figure S1. Performance of ROTS on DNA methylation sequencing data of human embryonic stem cells. (a) Overlaps of all statistically significantly differentially methylated 1 bp sites detected by MethylKit, RnBeads and ROTS at false discovery rate of 0.05. (b) Overlaps of 100 top-ranked sites detected by MethylKit, RnBeads and ROTS.

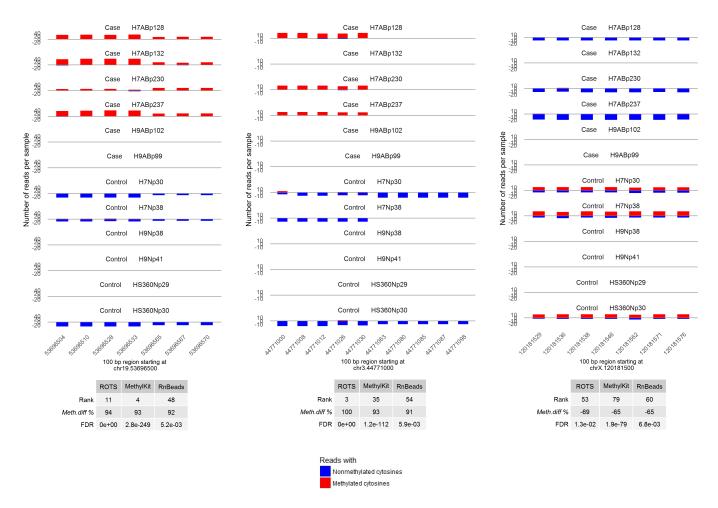


Figure S2. Representative examples of the 100 top differentially methylated 100 bp regions detected by MethylKit, RnBeads and ROTS presented in the RRBS human stem cell dataset. The 12 samples are shown as separate tracks. The numbers of mapped reads are denoted as stacked histograms; the size of the red bars reflects the number of nonconverted cytosines, implying methylation, while the blue bars denote bisulfite converted cytosines. CpG sites with less than 10 reads coverage are not shown.

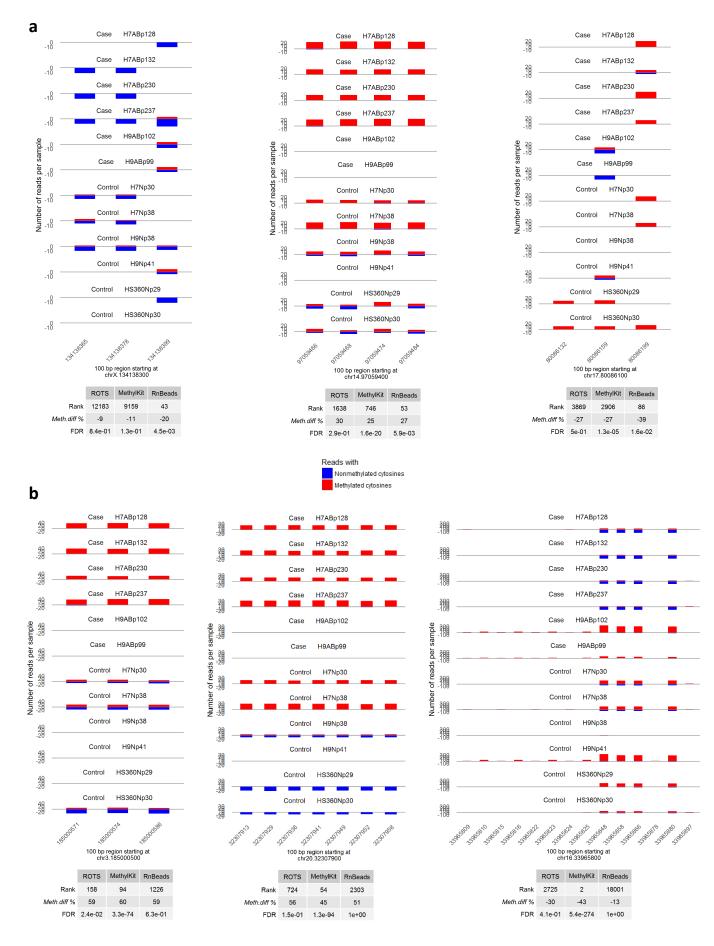


Figure S3. Representative examples of differentially methylated 100 bp regions detected among top 100 only by (a) RnBeads, or (b) MethylKit. The 12 samples are shown as separate tracks. The numbers of mapped reads are denoted as stacked histograms.

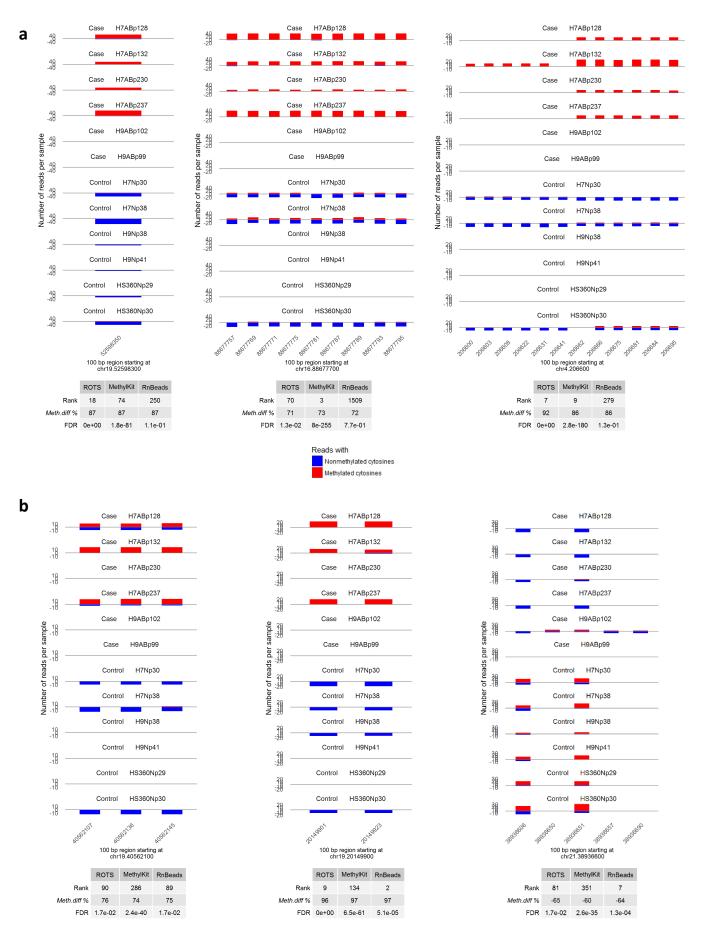


Figure S4. Representative examples of differentially methylated 100 bp regions detected among top 100 only by **(a)** ROTS and MethylKit, or **(b)** ROTS and RnBeads. The 12 samples are shown as separate tracks. The numbers of mapped reads are denoted as stacked histograms.

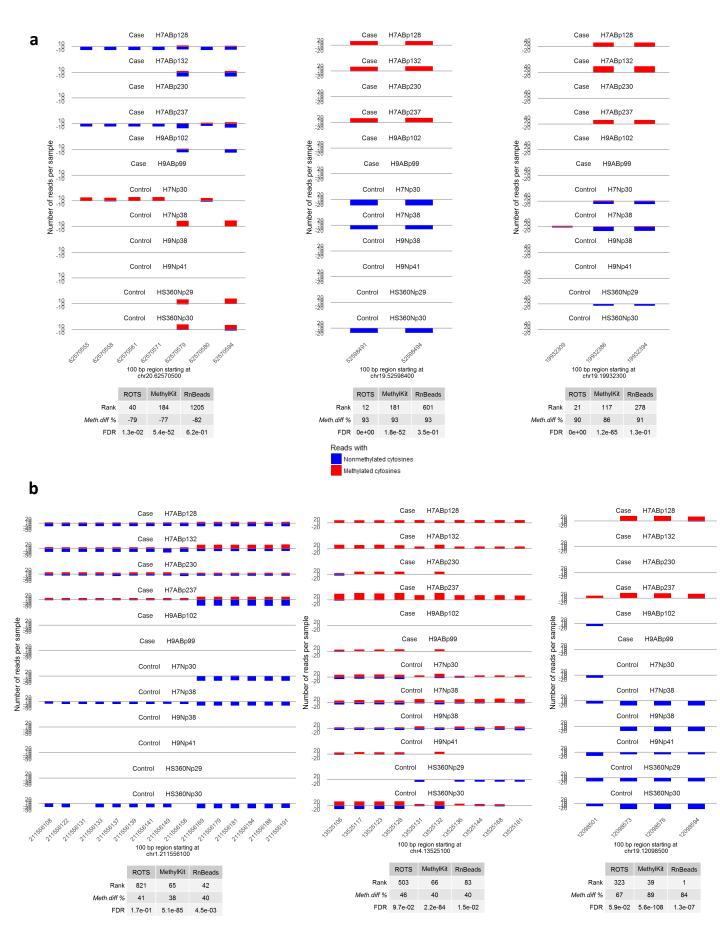
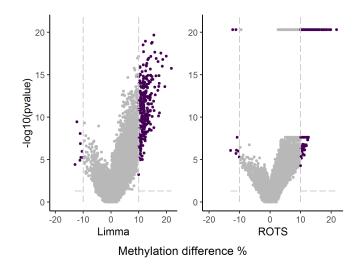


Figure S5. Representative examples of differentially methylated 100 bp regions detected among top 100 only by (a) ROTS, or (b) MethylKit and RnBeads. The 12 samples are shown as separate tracks. The numbers of mapped reads are denoted as stacked histograms.



а

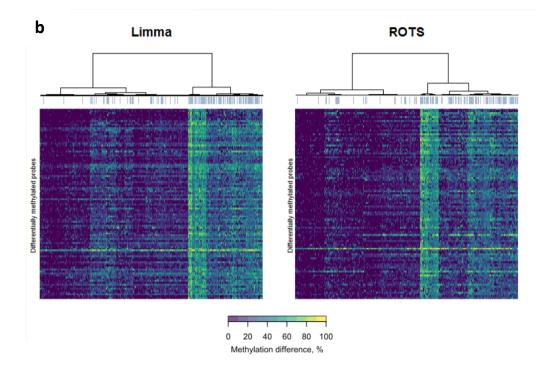
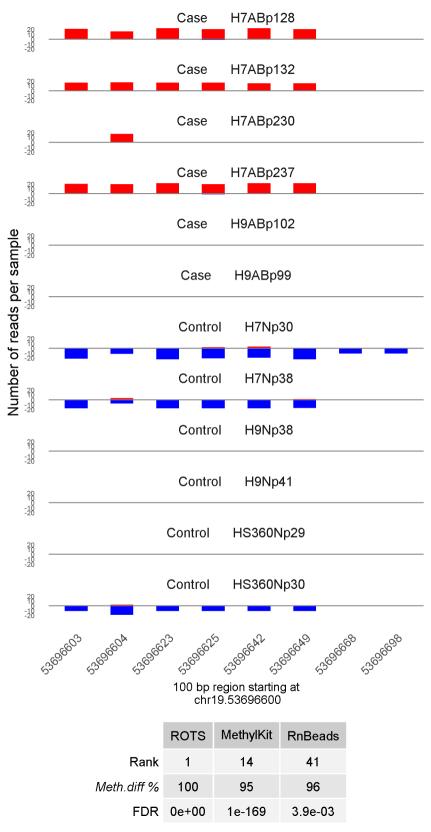
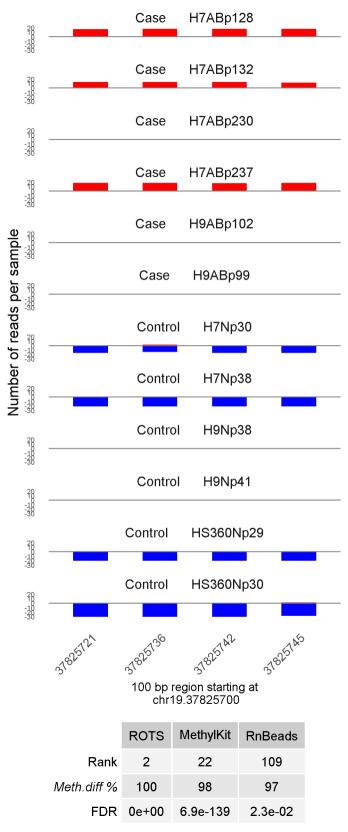
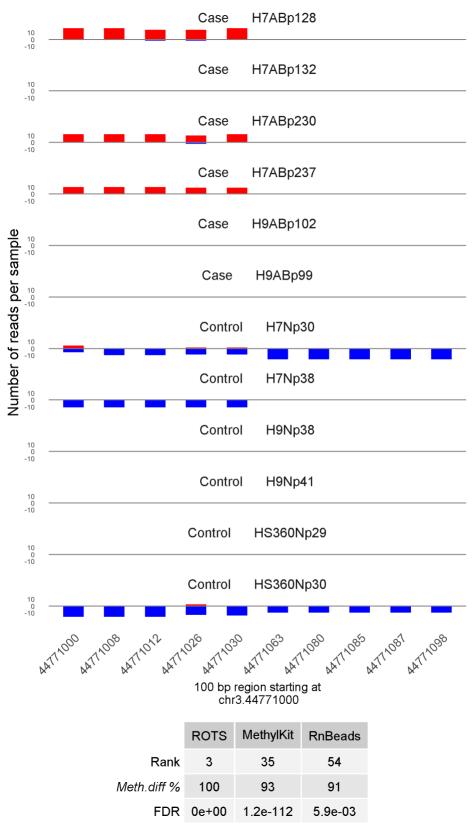
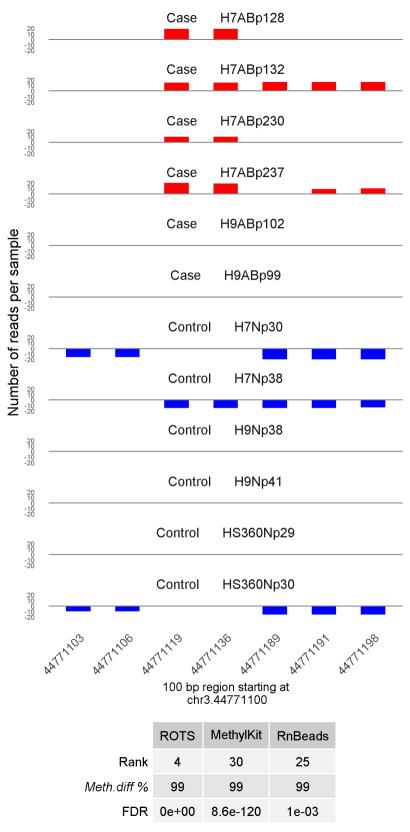


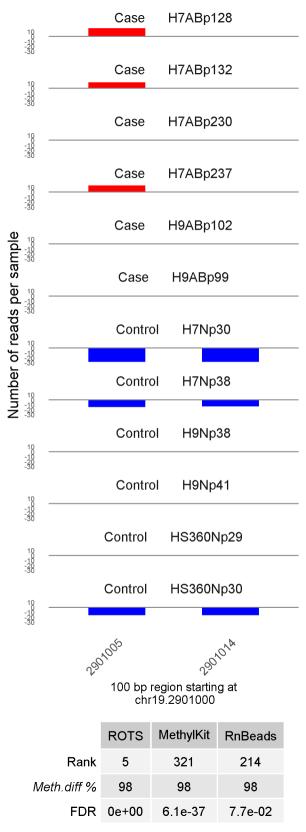
Figure S6. Performance of ROTS on the array-based DNA methylation study of clear cell renal cell carcinoma (ccRCC) patients. **(a)** Volcano plots of methylation difference levels (x-axis) versus their significance p-values (y-axis) for ROTS and Limma. The purple dots represent 296 differentially methylated probes with p-value < 0.05 and methylation change > 10%. **(b)** Hierarchical clustering of the metastatic (n = 132) and non-metastatic (n = 266) ccRCC patients (columns) across the top-ranking 100 differentially methylated probes (rows) with ROTS and Limma. The colors indicate the methylation level of each probe: yellow is high, green is intermediate, and blue is low methylation level. Metastatic patient samples are marked in blue in the dendrogram, non-metastatic in white.

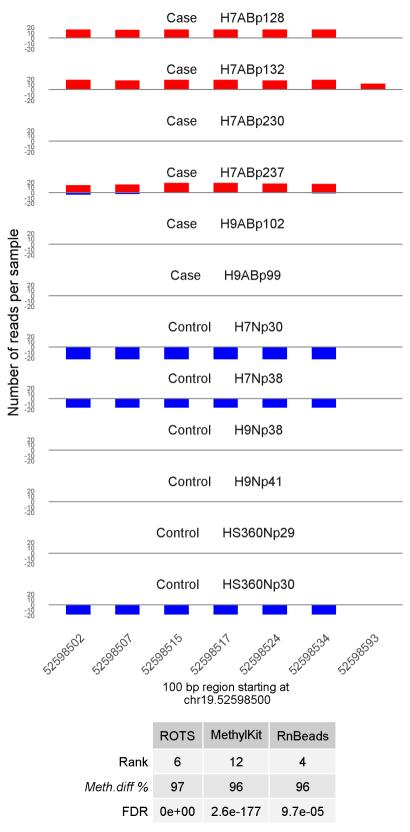


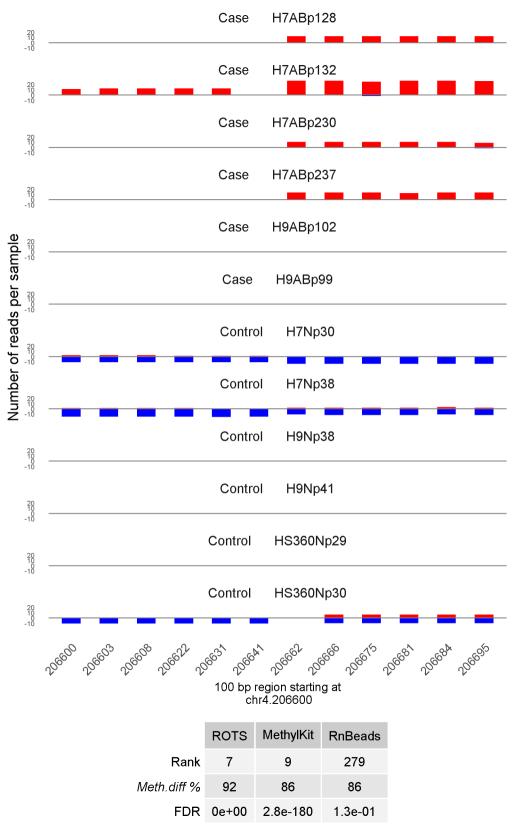


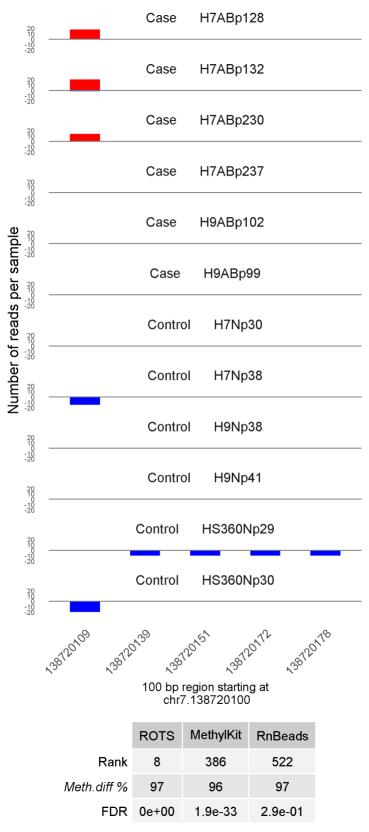


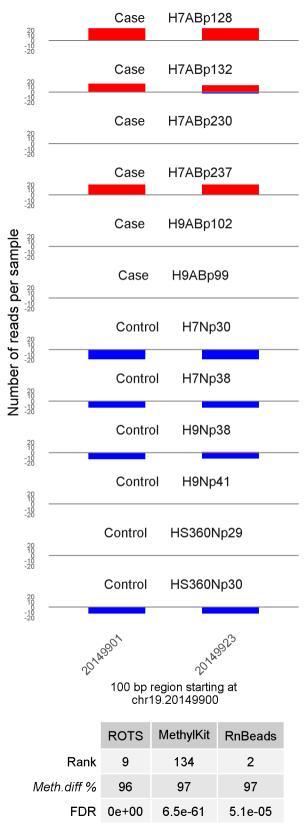


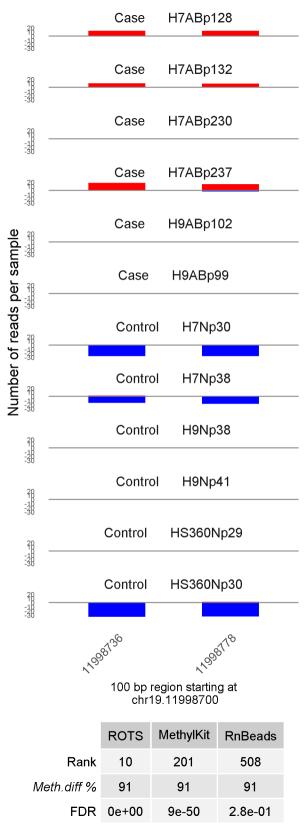


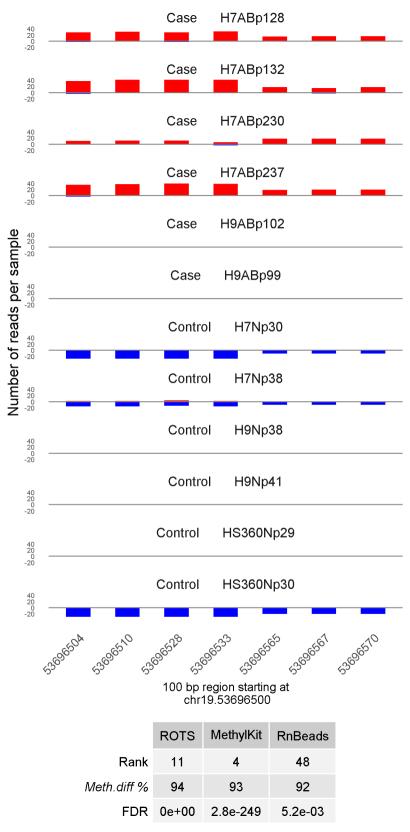


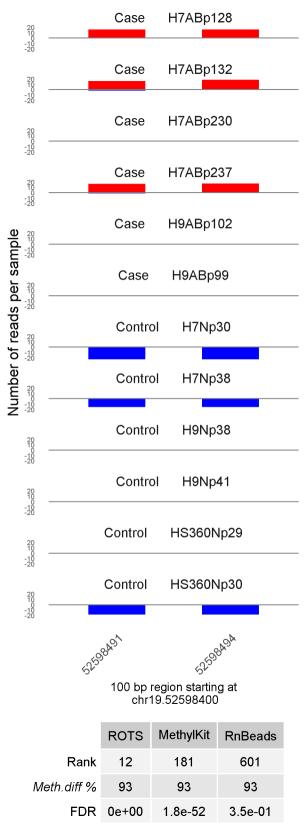


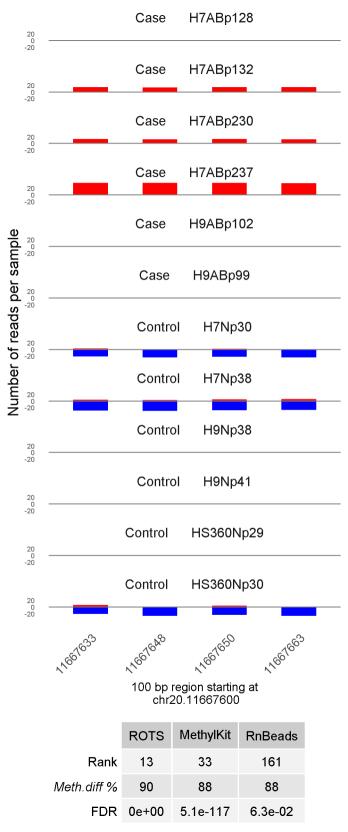


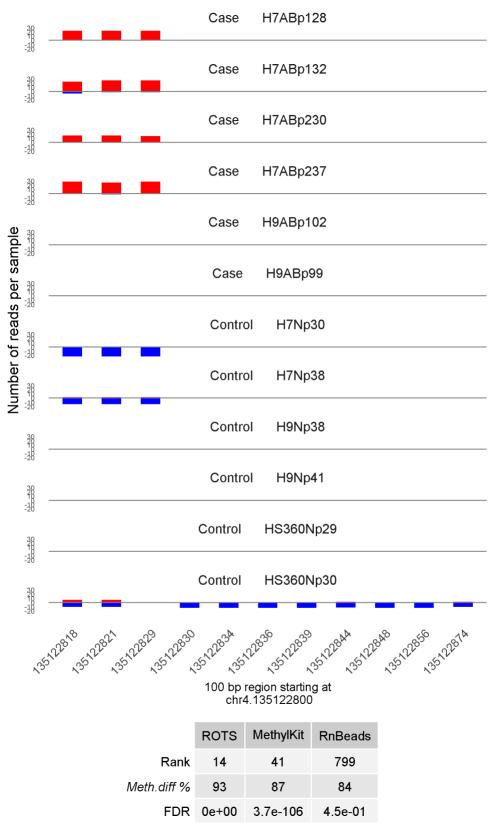


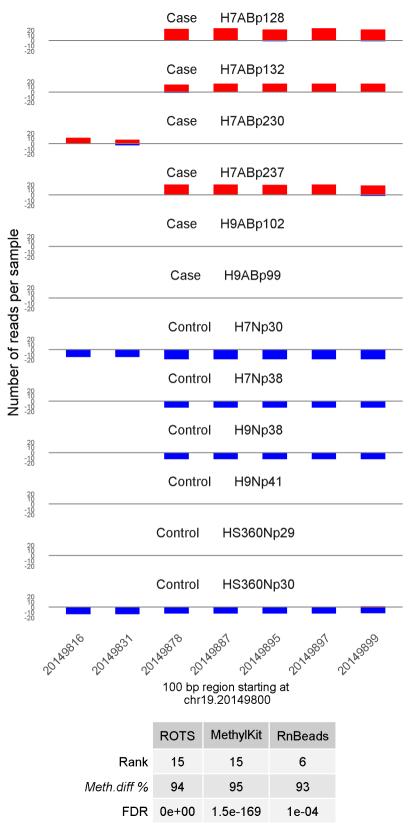


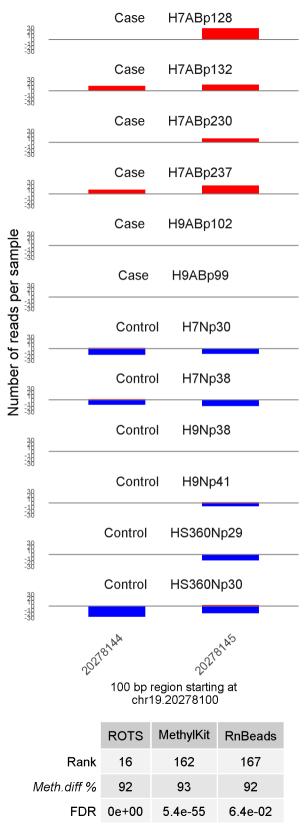


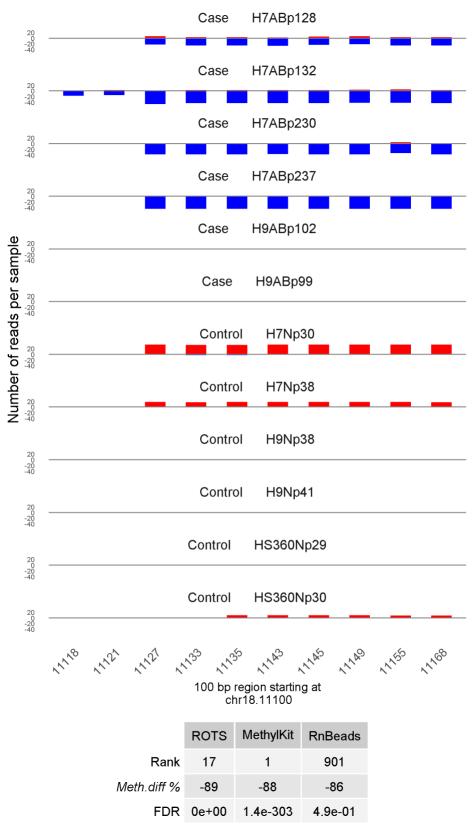


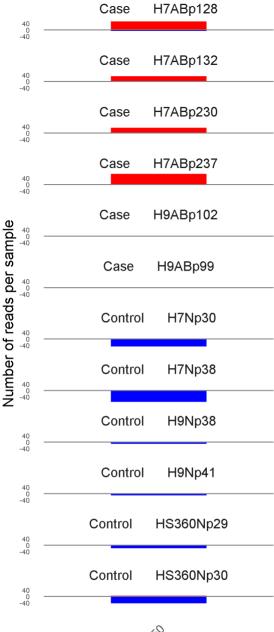








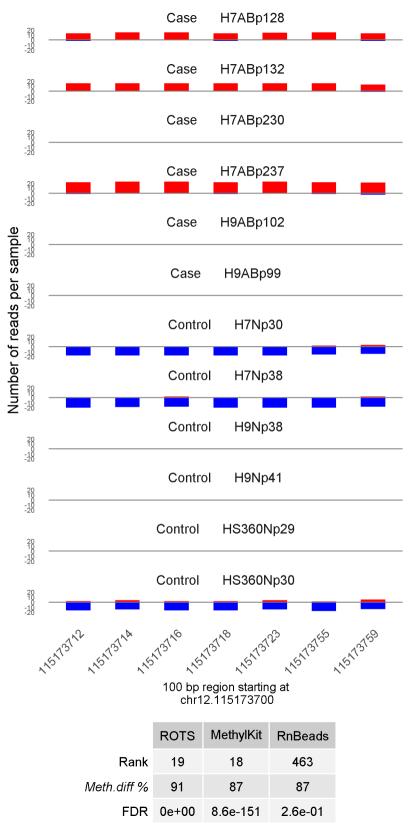


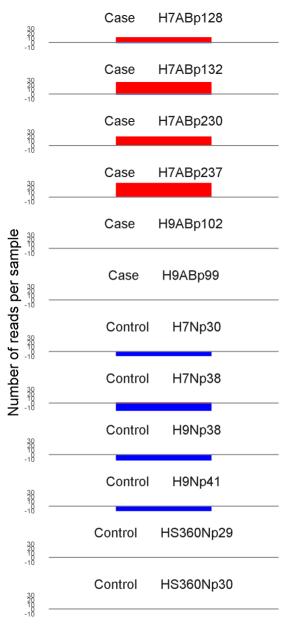




100 bp region starting at chr19.52598300

	ROTS	MethylKit	RnBeads
Rank	18	74	250
Meth.diff %	87	87	87
FDR	0e+00	1.8e-81	1.1e-01

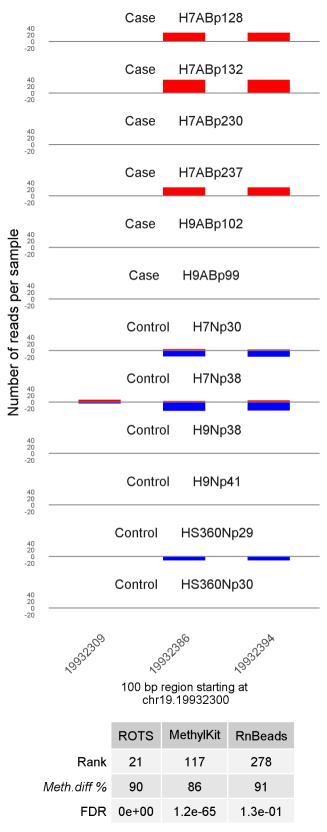


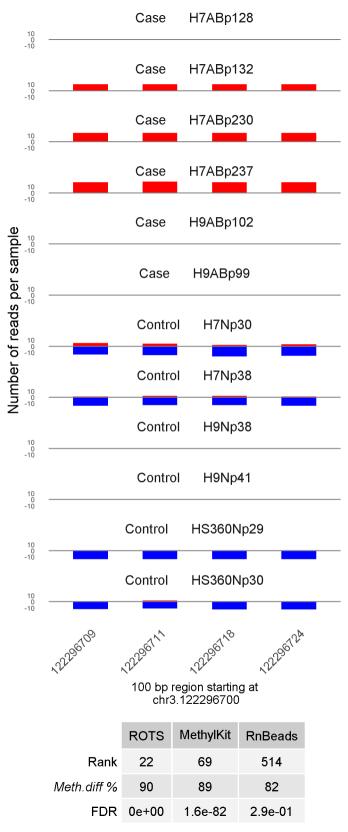


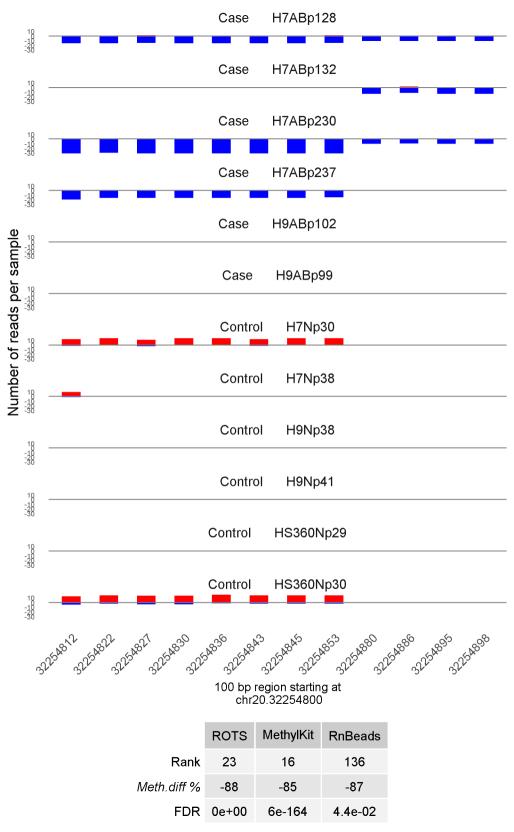


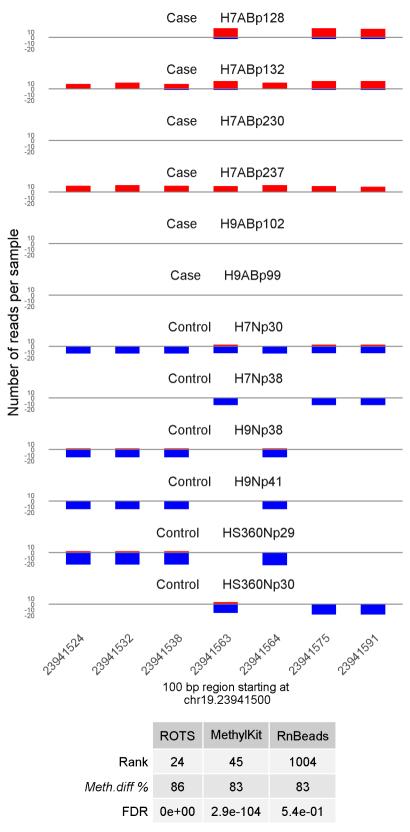
100 bp region starting at chr1.74663700

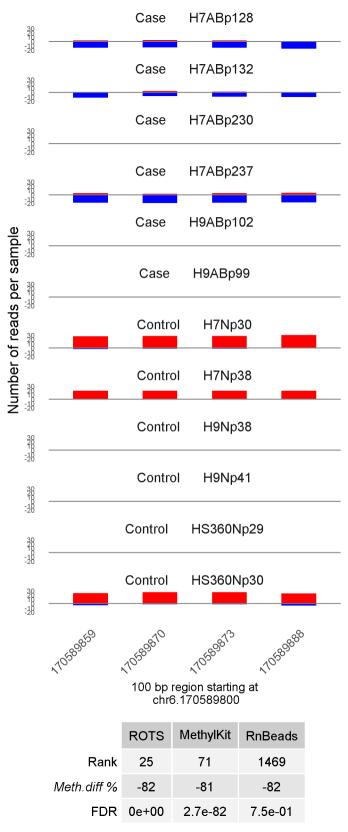
	ROTS	MethylKit	RnBeads
Rank	20	475	1208
Meth.diff %	89	91	88
FDR	0e+00	6.1e-29	6.3e-01

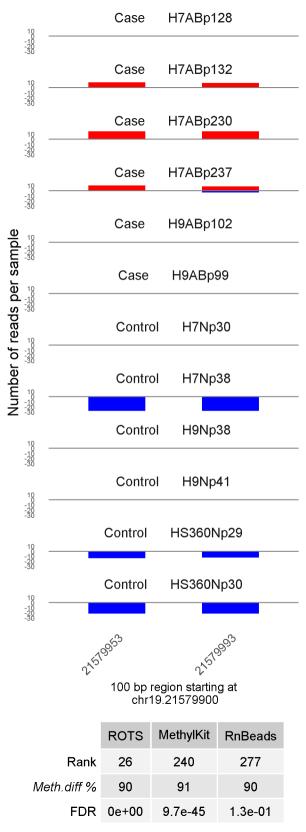


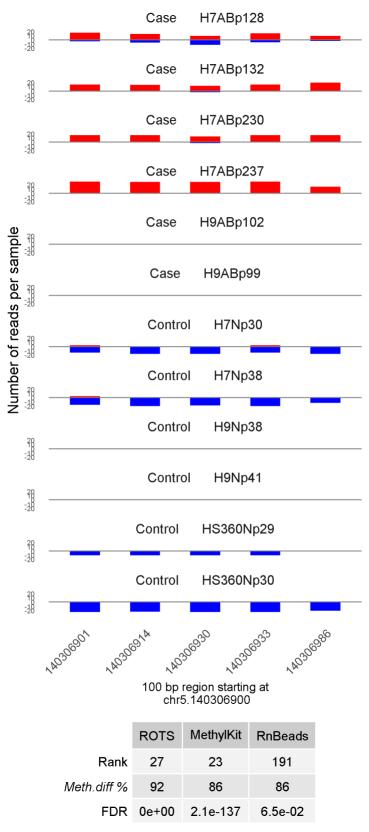


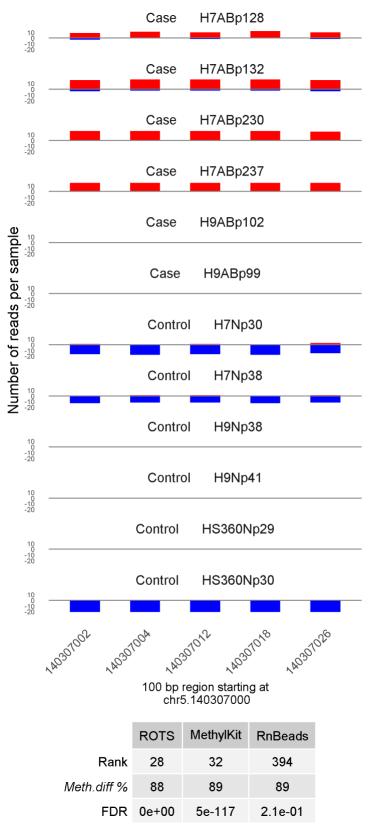


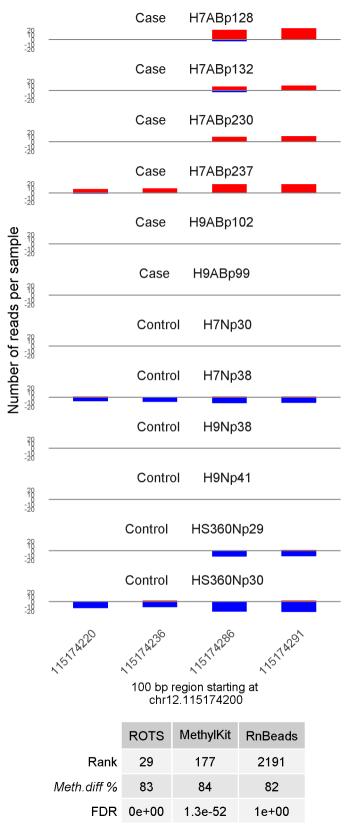


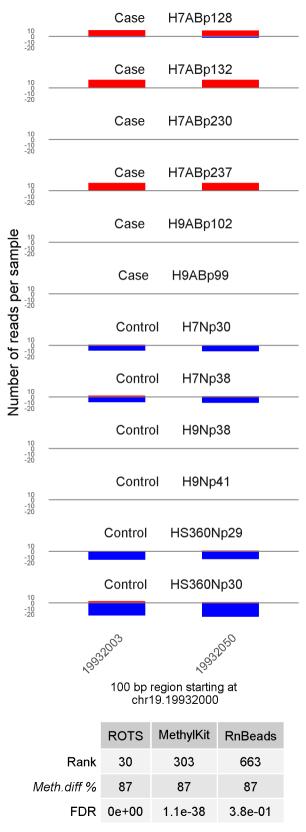


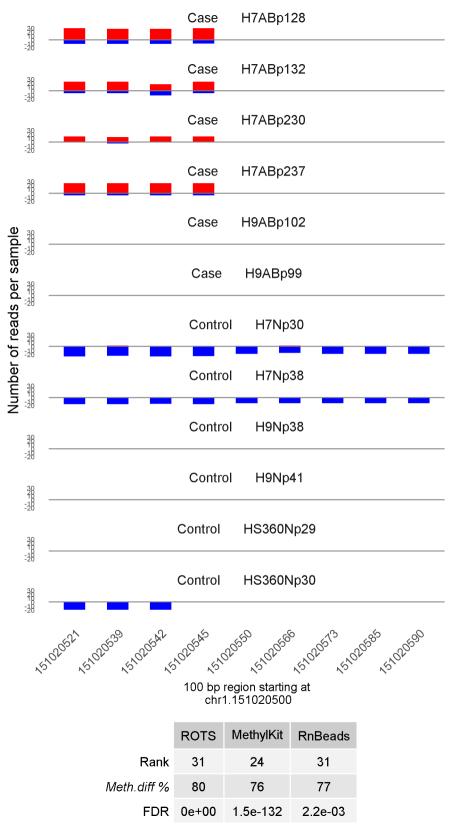


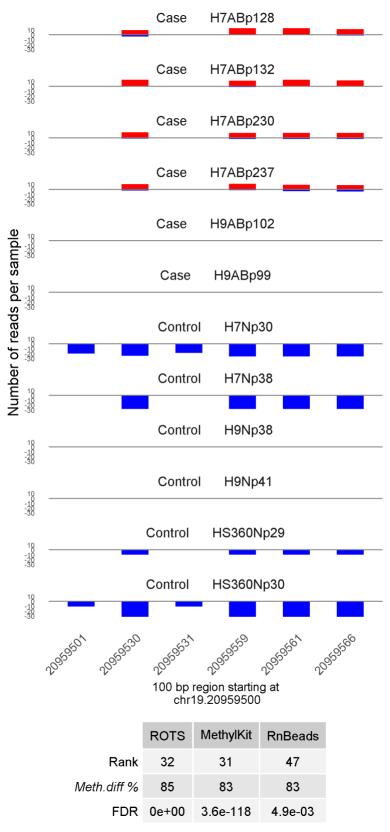


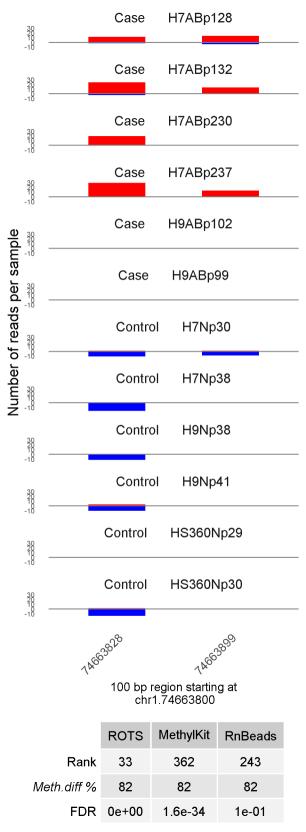


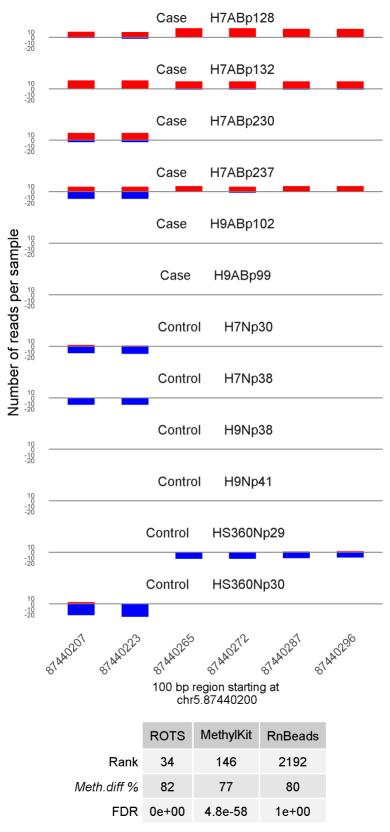


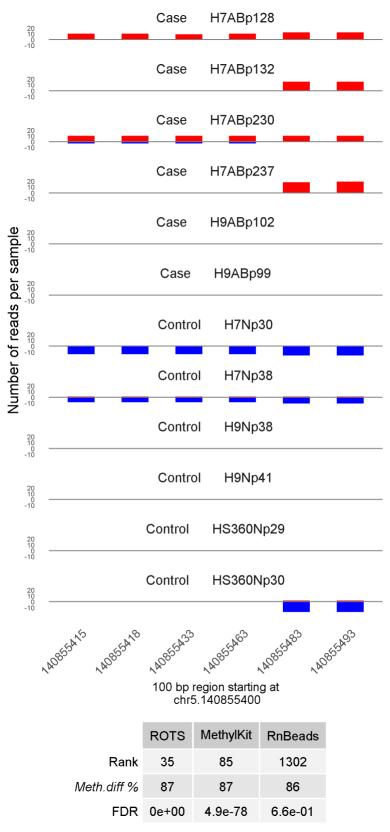


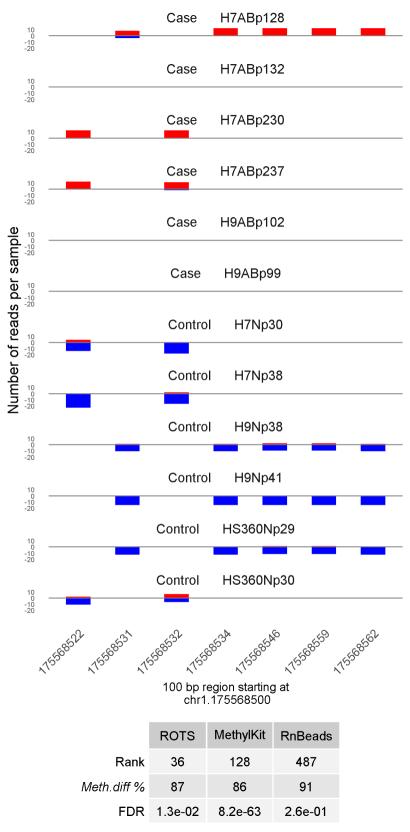


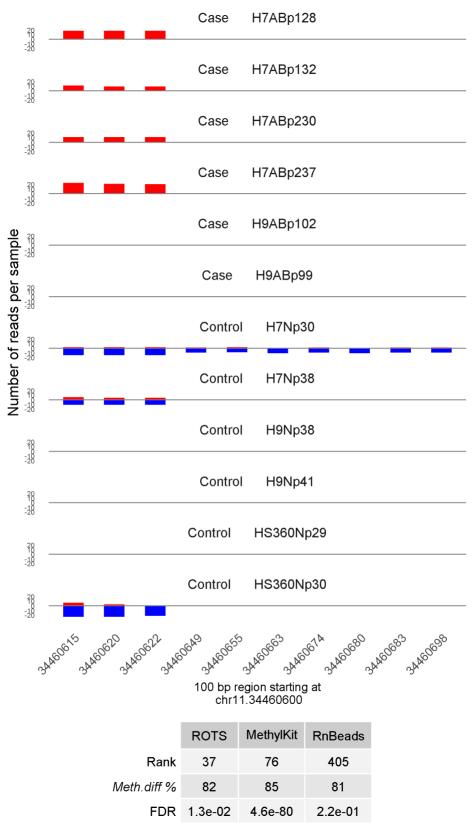


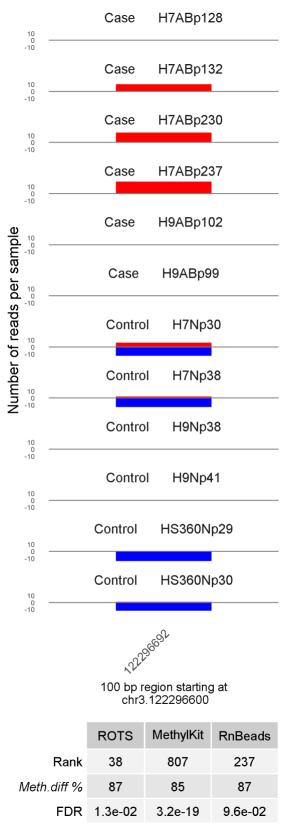


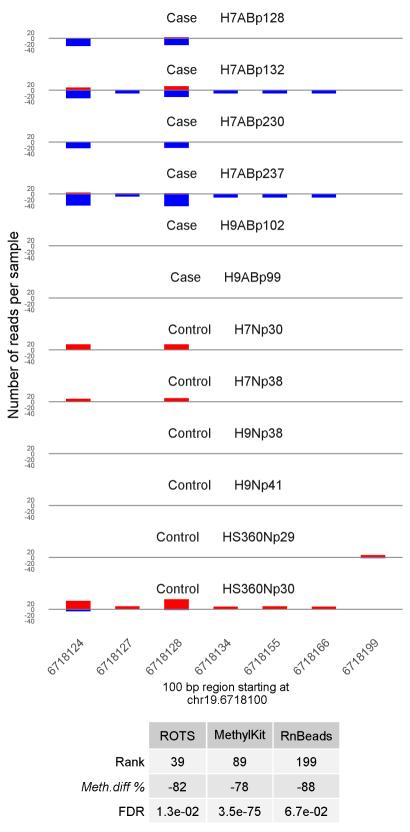


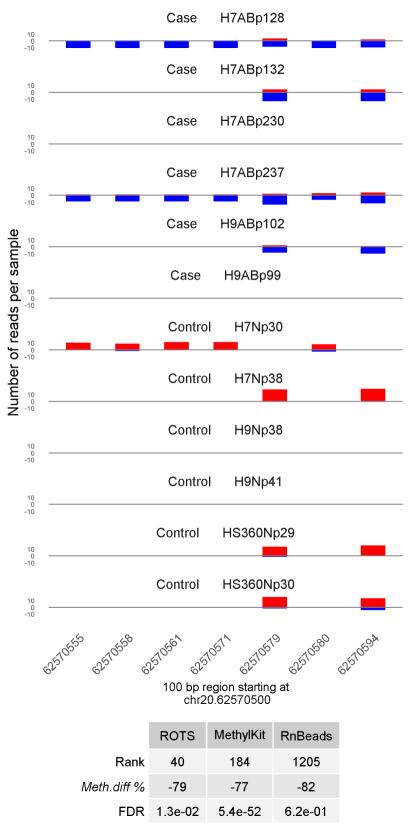


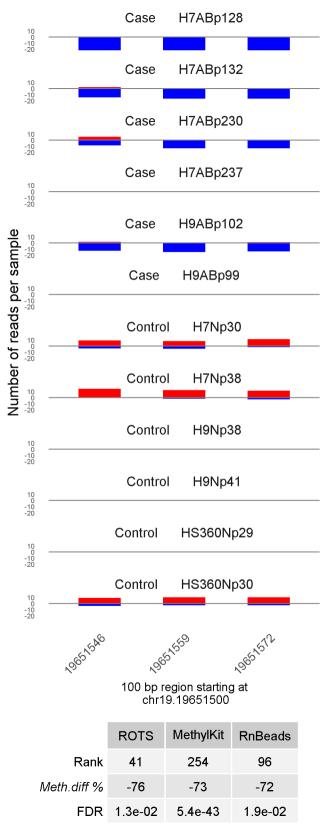


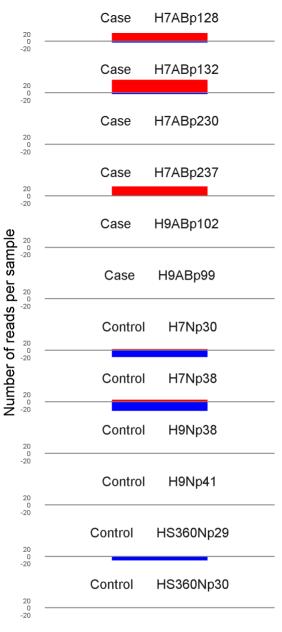








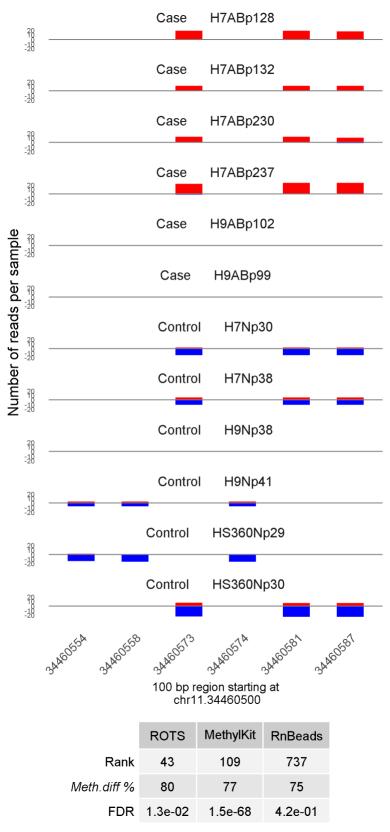


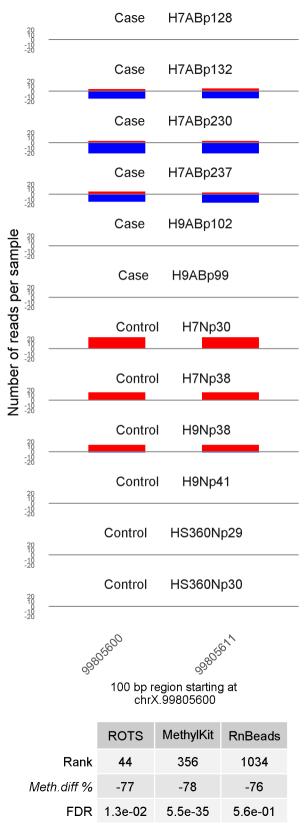


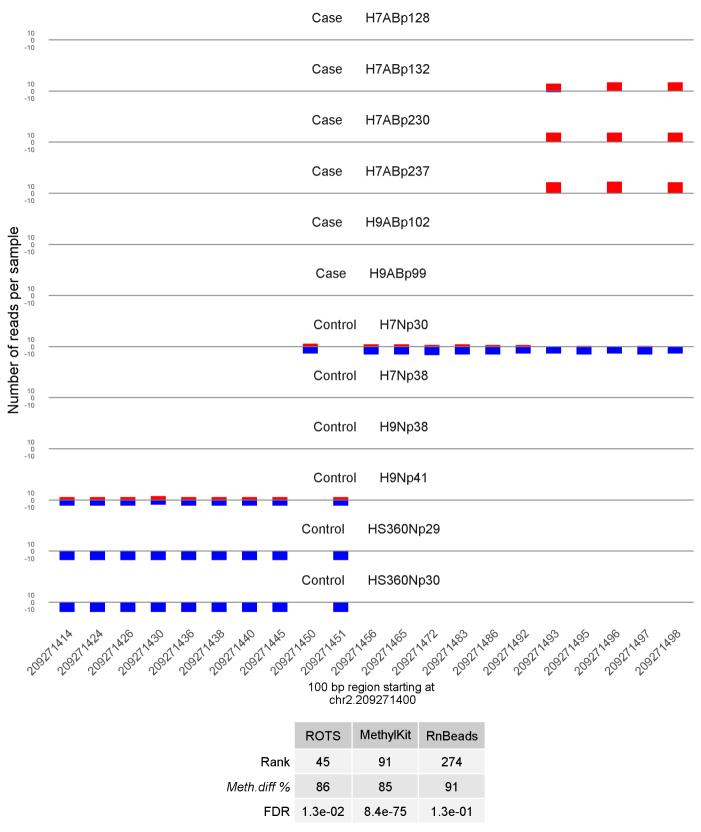


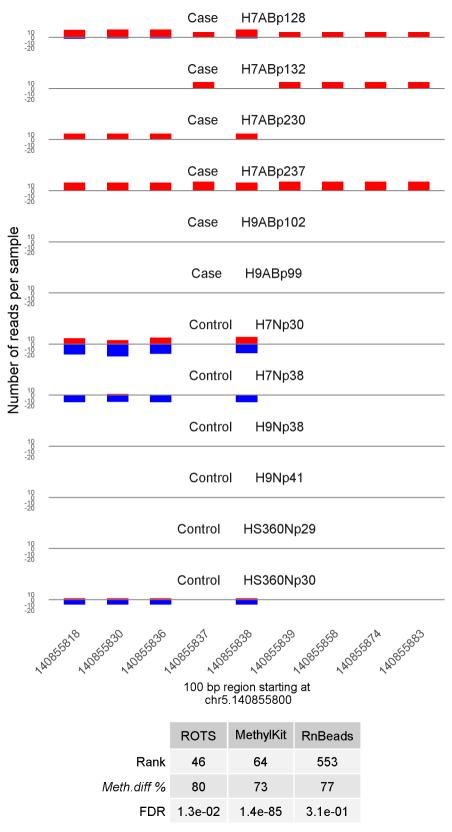
100 bp region starting at chr19.19932400

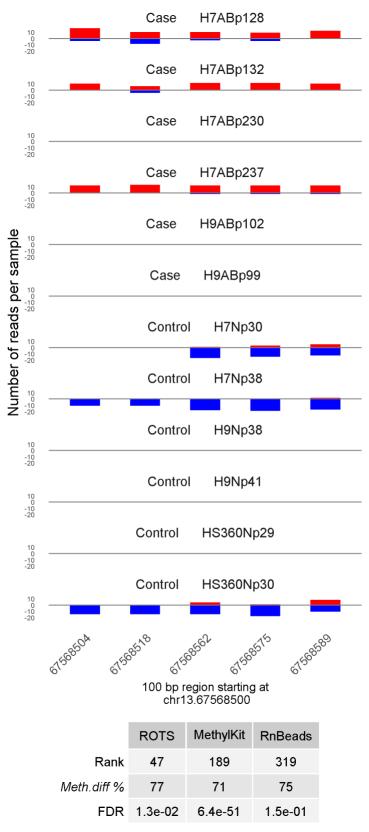
	ROTS	MethylKit	RnBeads
Rank	42	651	2193
Meth.diff %	81	79	81
FDR	1.3e-02	9.6e-23	1e+00

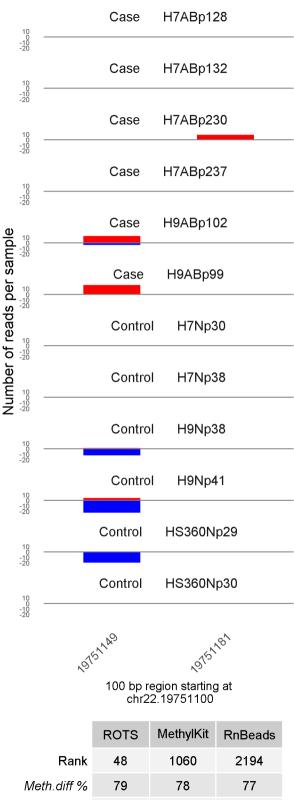






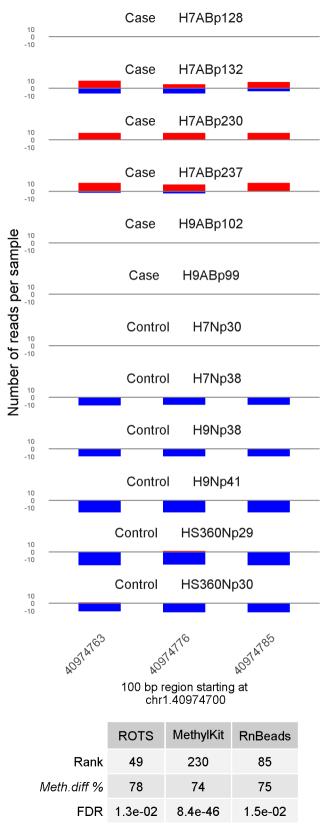


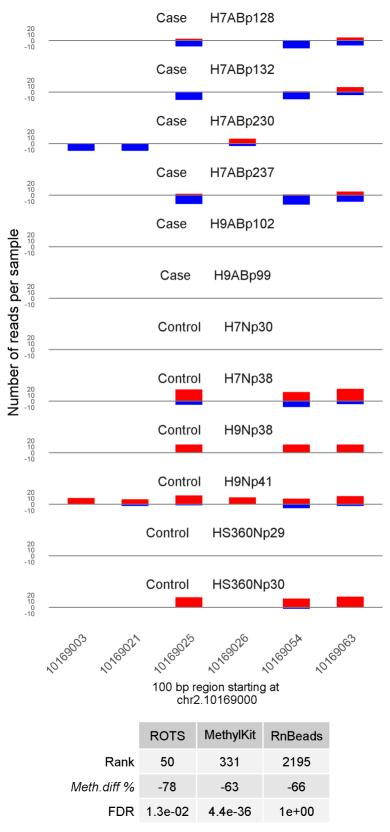


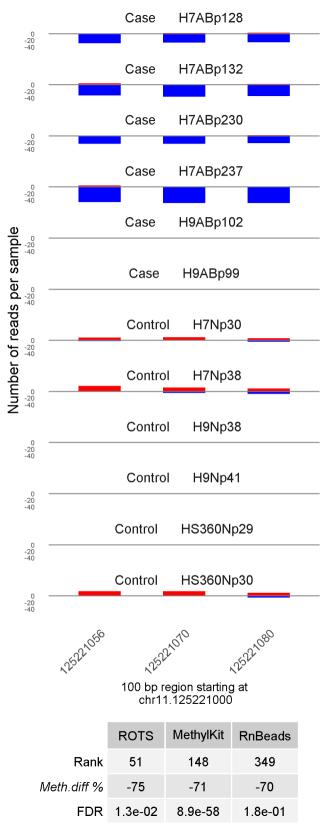


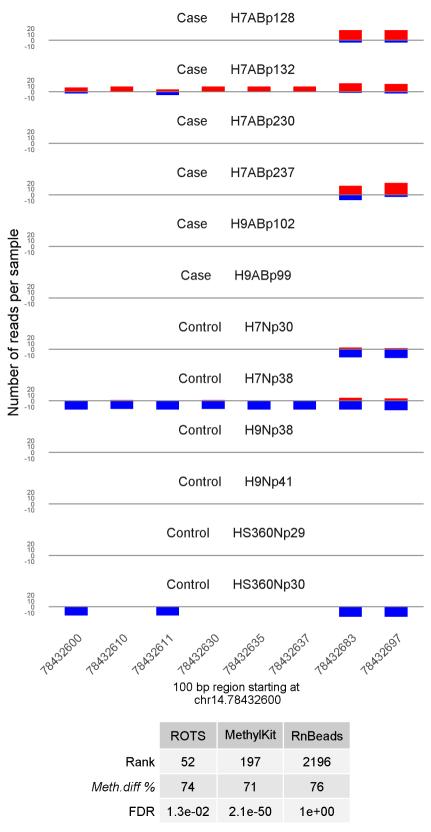
FDR 1.3e-02 3.2e-15

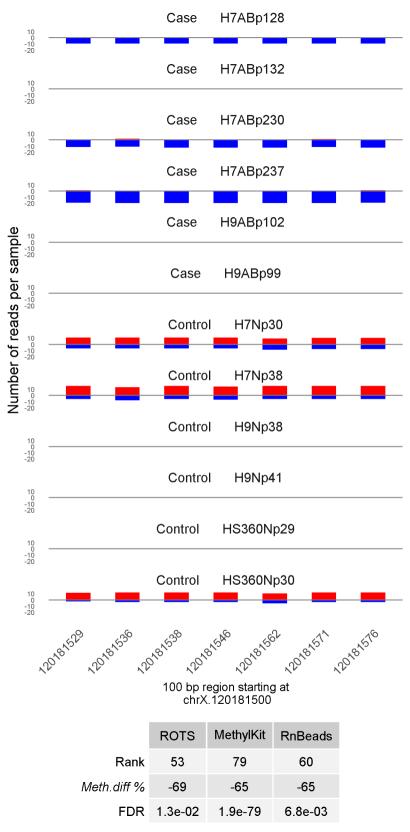
1e+00

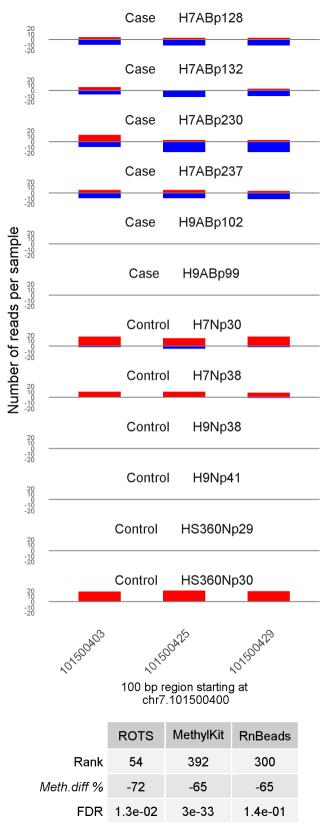


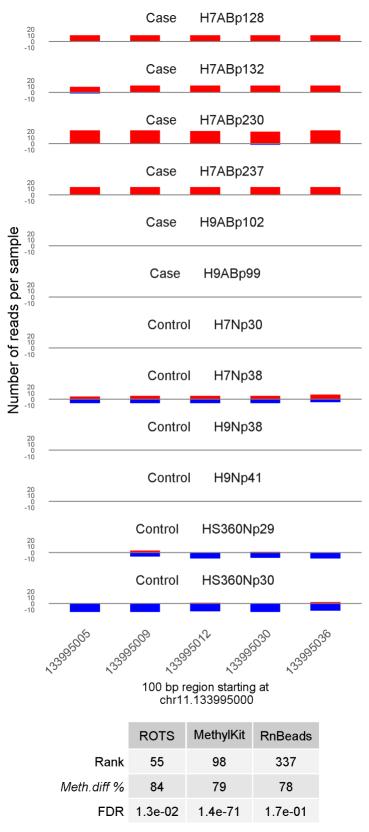


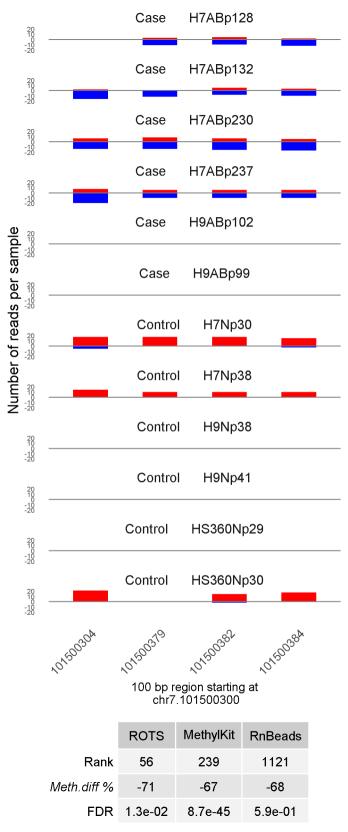


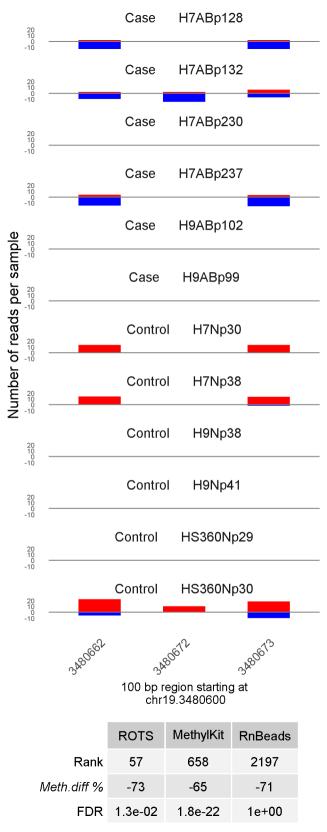


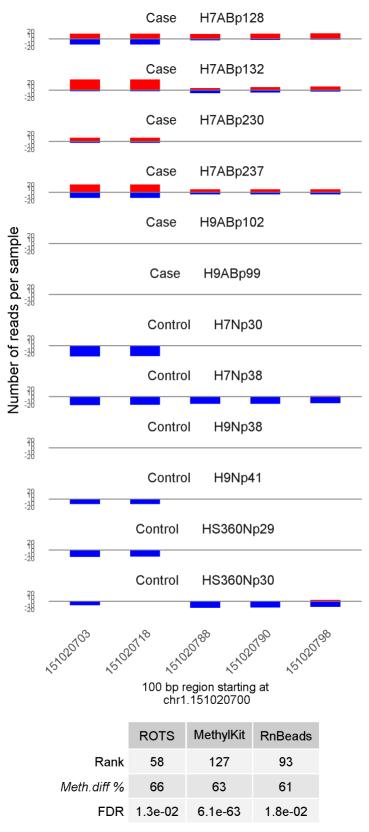


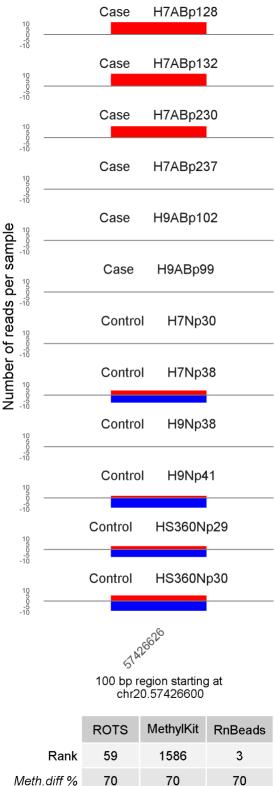






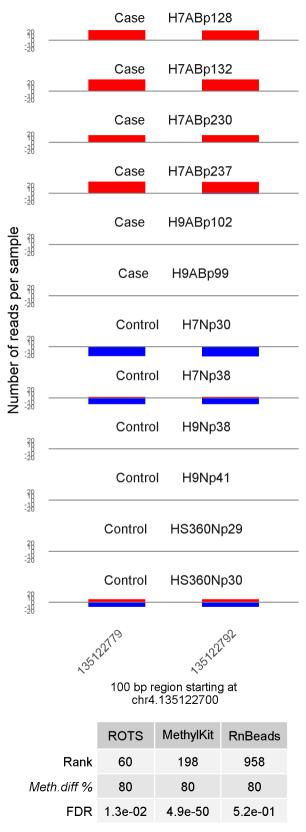


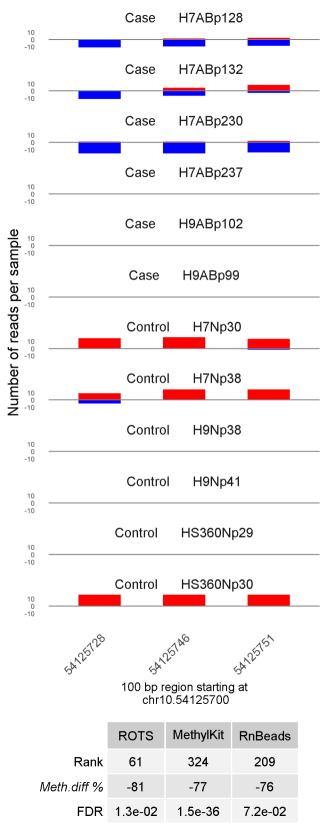


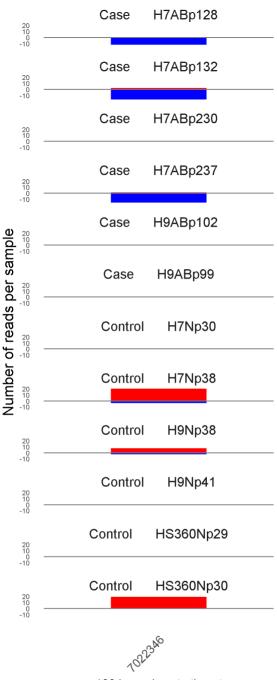


FDR 1.3e-02 1.9e-10

6.3e-05

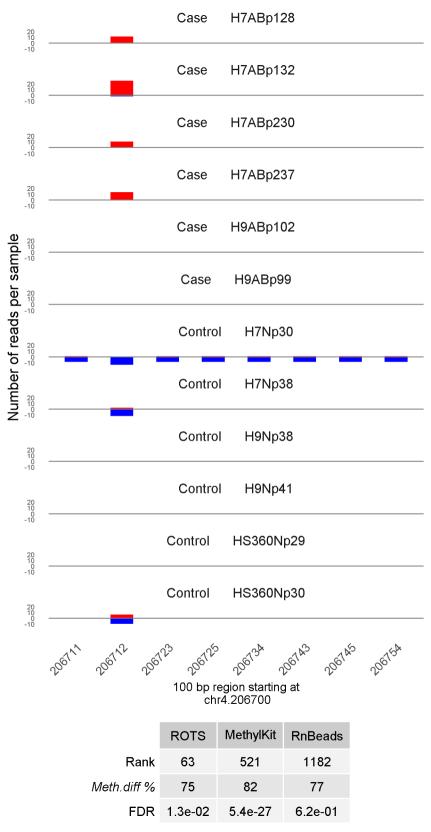


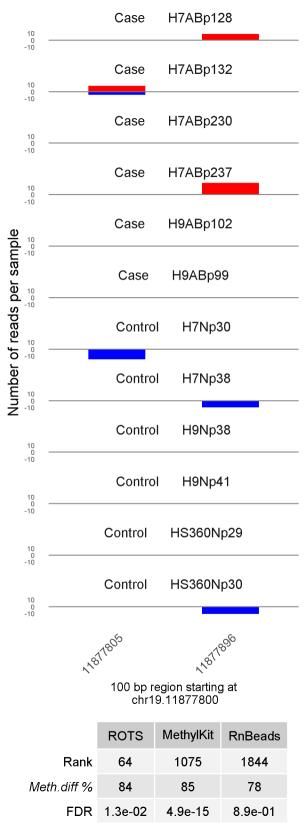


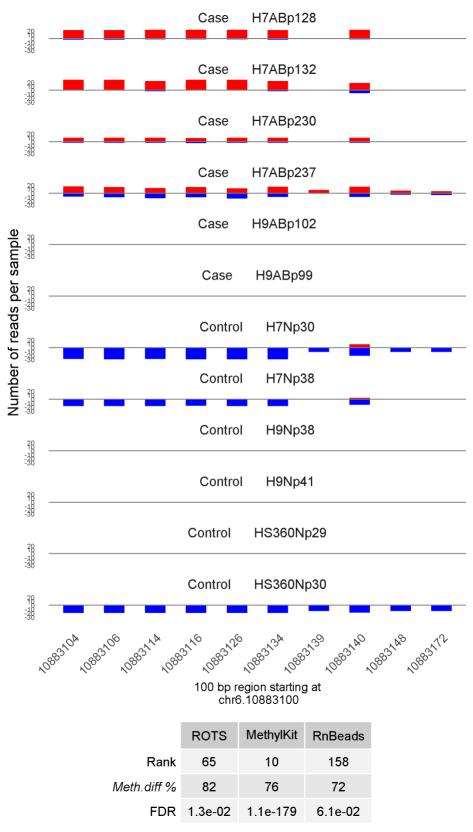


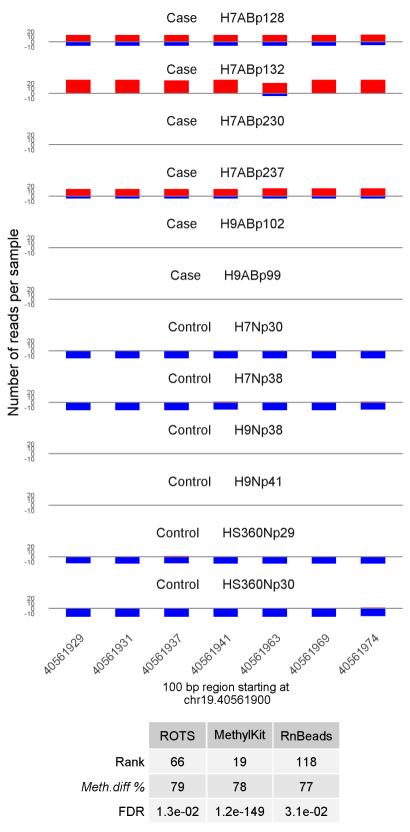
100 bp region starting at chr5.7022300

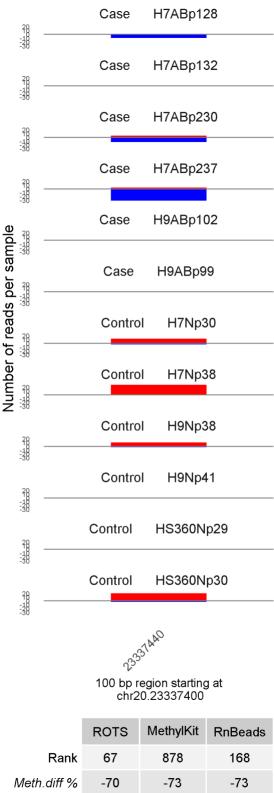
	ROTS	MethylKit	RnBeads
Rank	62	991	2198
Meth.diff %	-79	-81	-86
FDR	1.3e-02	3.4e-16	1e+00



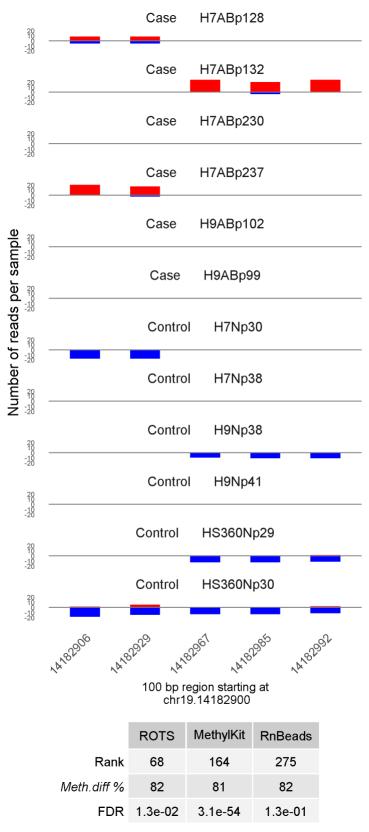


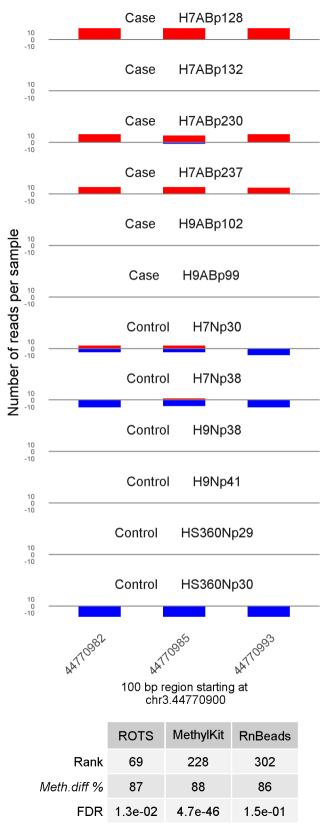


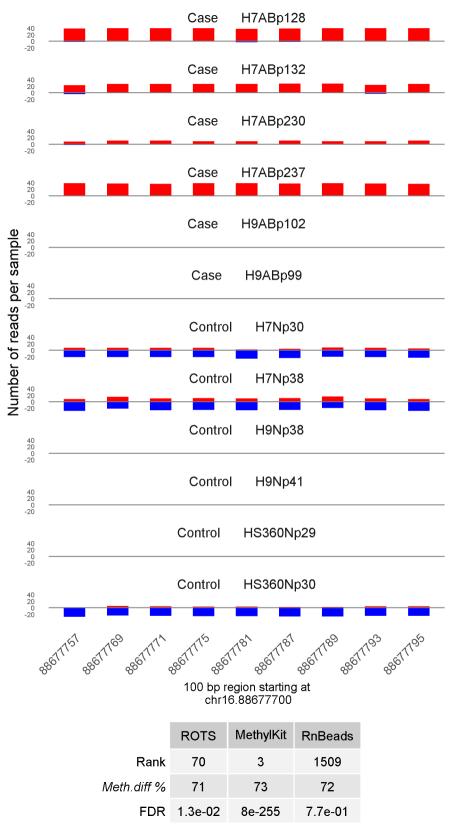


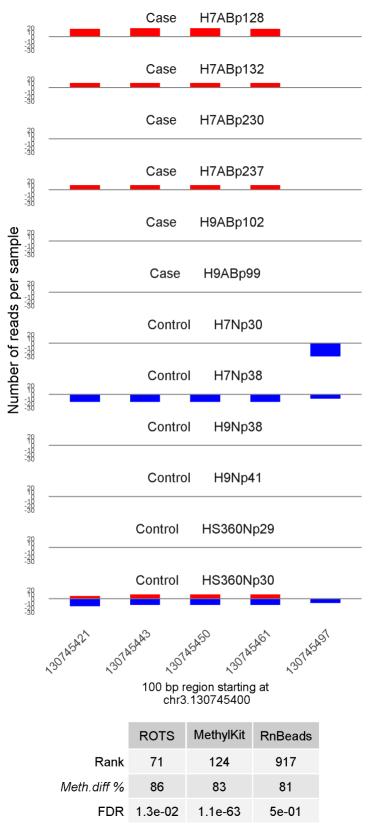


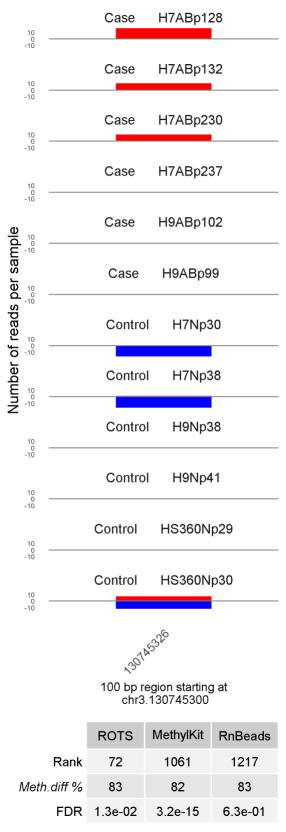
FDR 1.3e-02 5.8e-18 6.4e-02

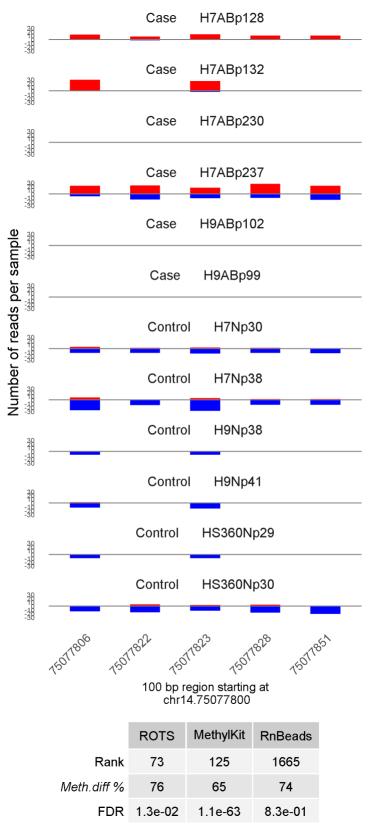


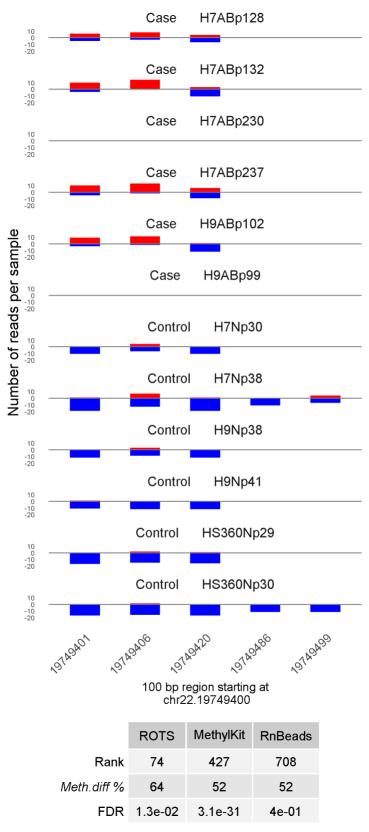


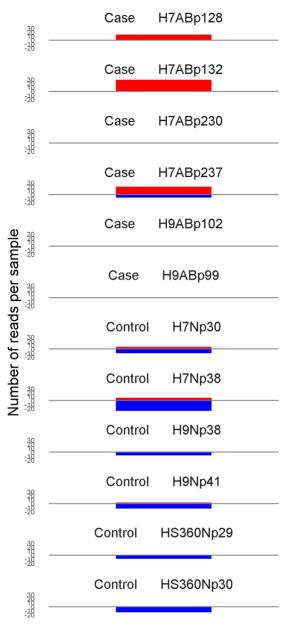








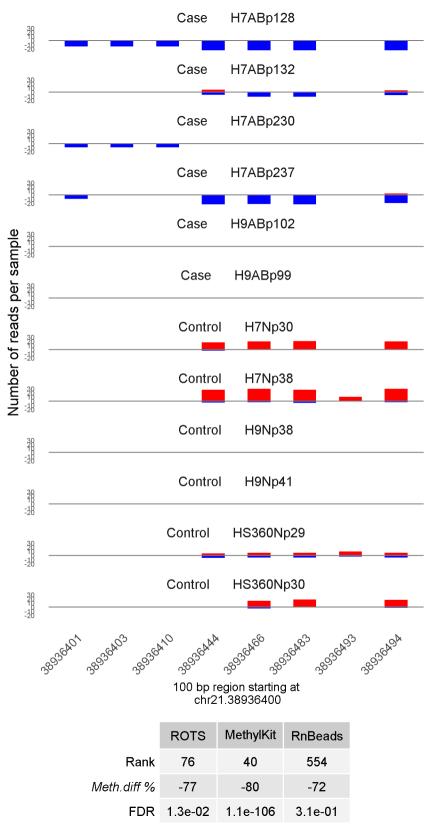


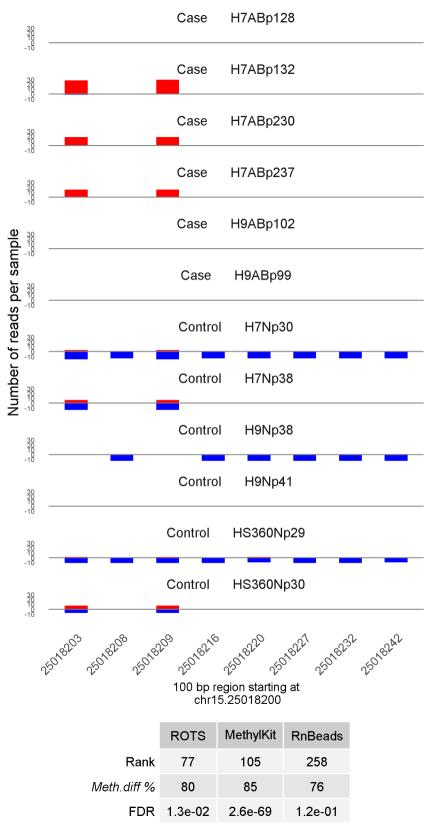


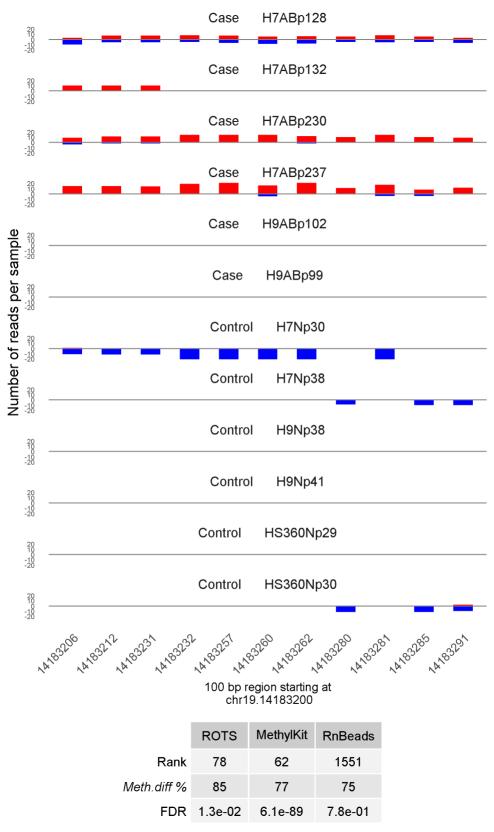


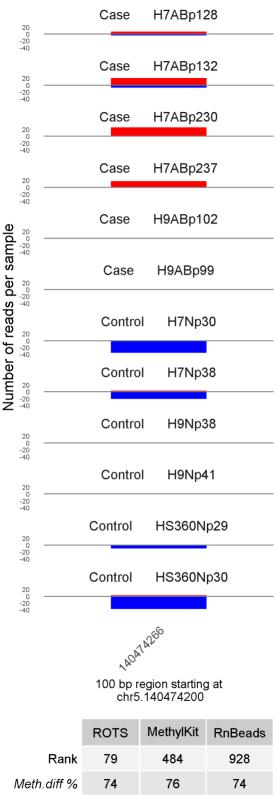
100 bp region starting at chr14.75077700

	ROTS	MethylKit	RnBeads
Rank	75	734	749
Meth.diff %	77	73	77
FDR	1.3e-02	8.8e-21	4.2e-01





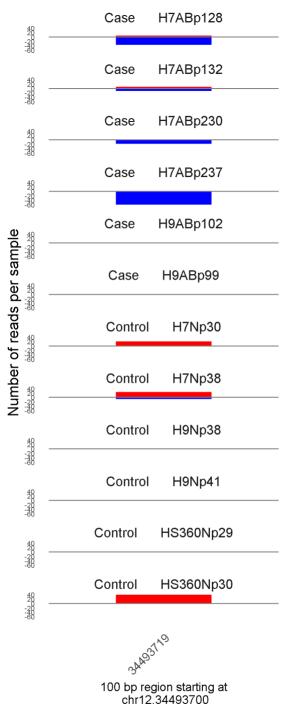




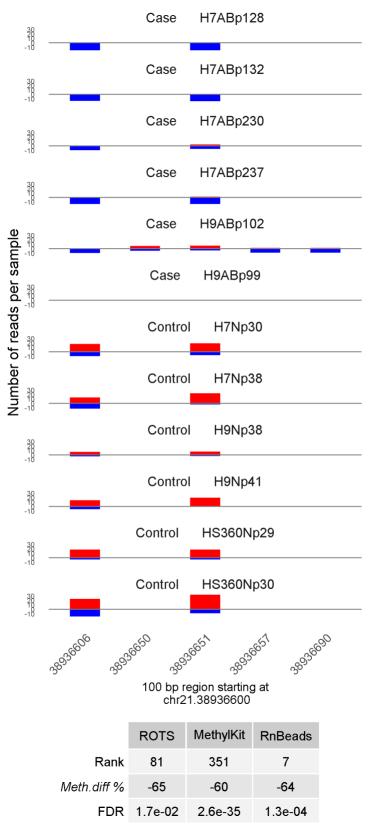
FDR 1.7e-02

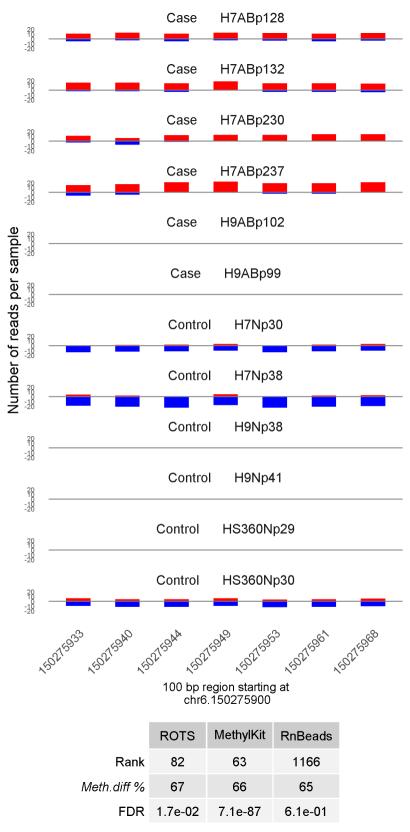
2e-28

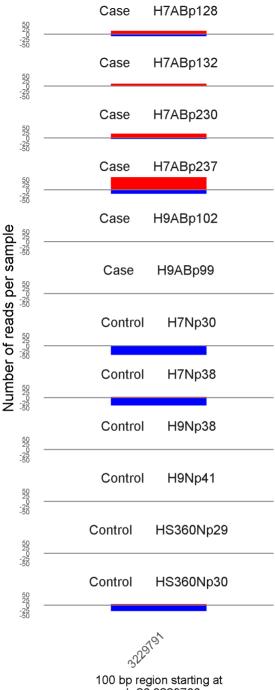
5e-01



	ROTS	MethylKit	RnBeads
Rank	80	325	1365
Meth.diff %	-79	-81	-79
FDR	1.7e-02	1.6e-36	7e-01

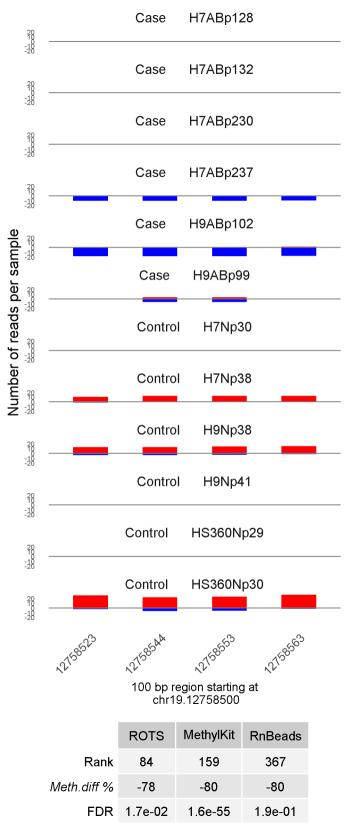


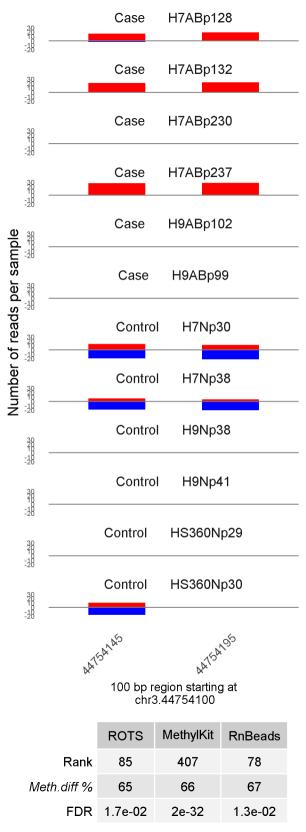


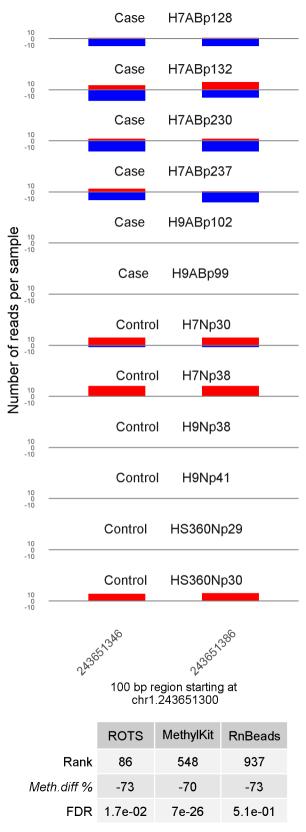


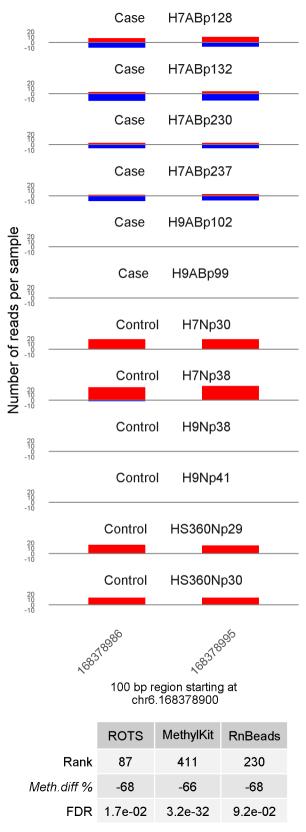
chr20.3229700

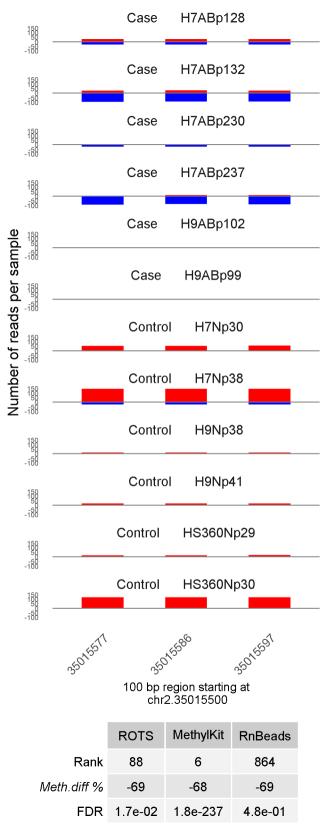
	ROTS	MethylKit	RnBeads
Rank	83	458	201
Meth.diff %	67	67	53
FDR	1.7e-02	7e-30	6.7e-02

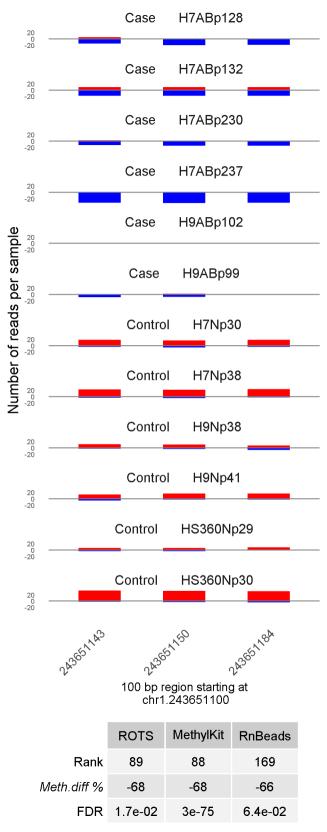


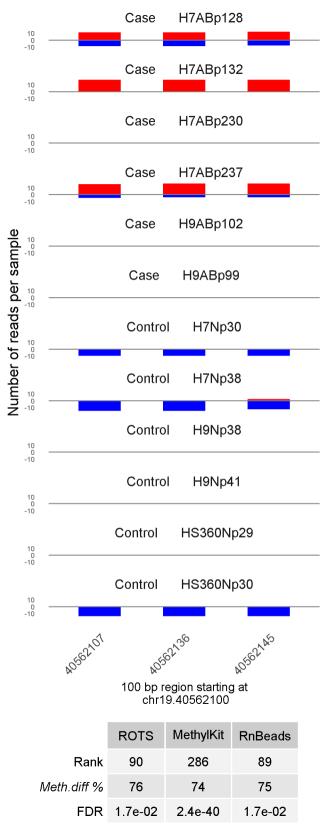


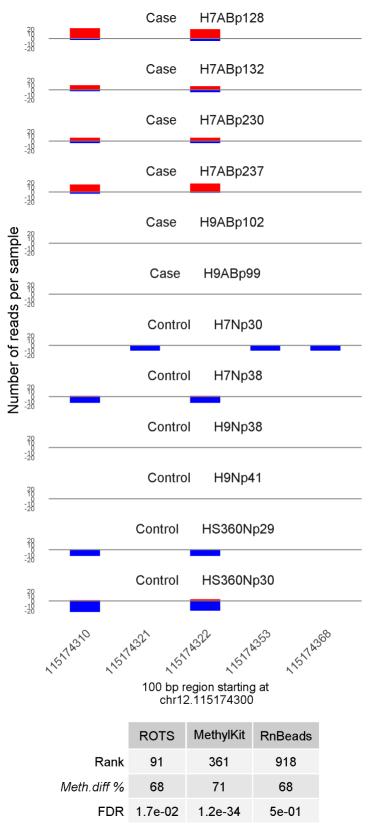


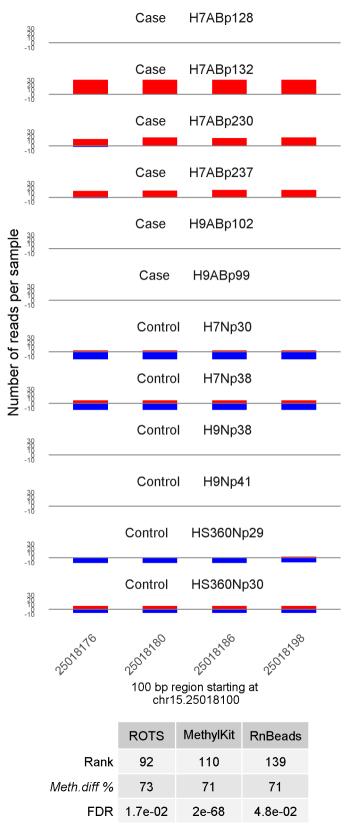


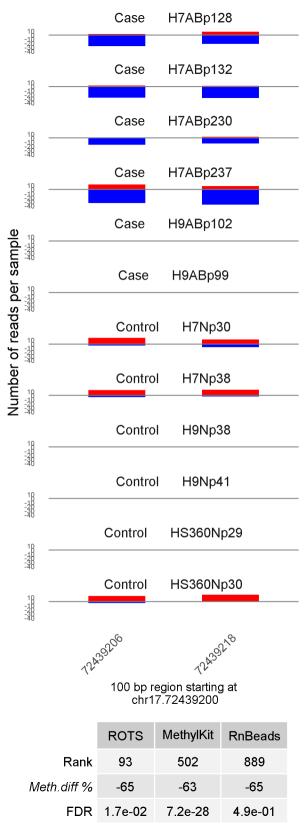


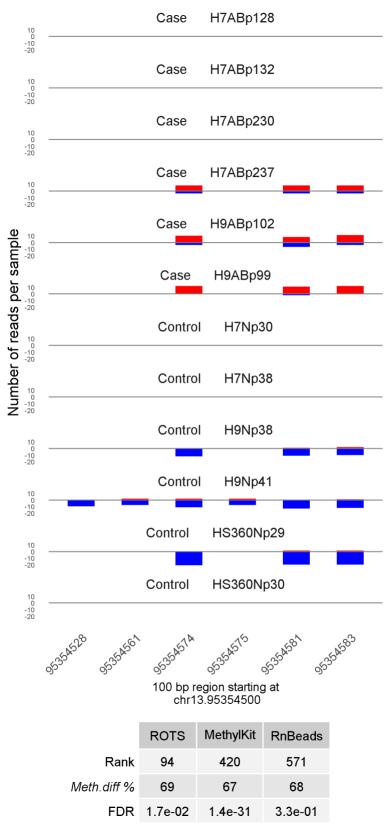


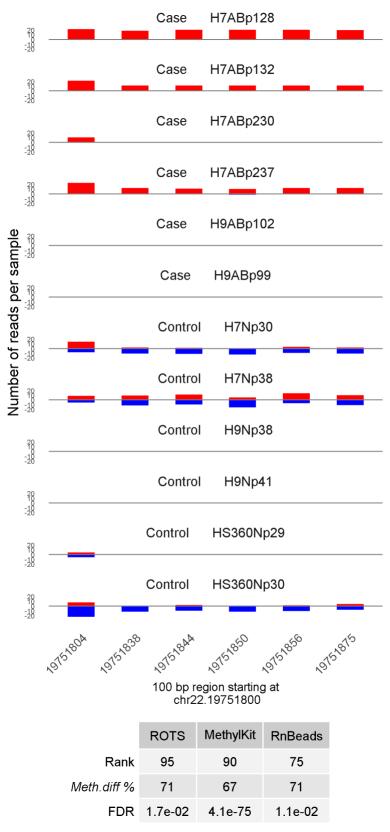


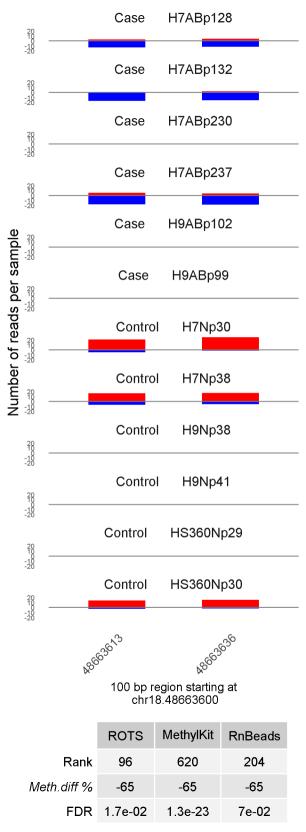


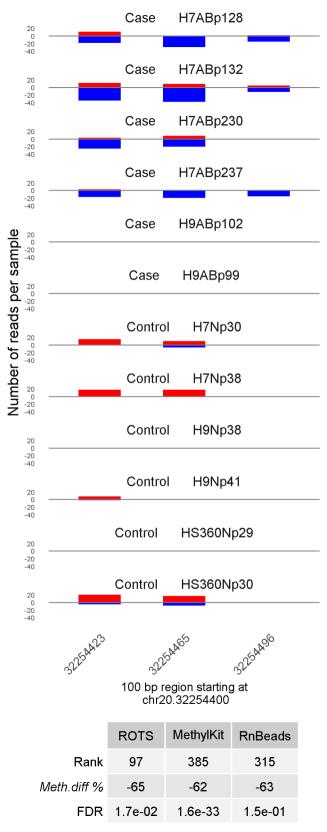


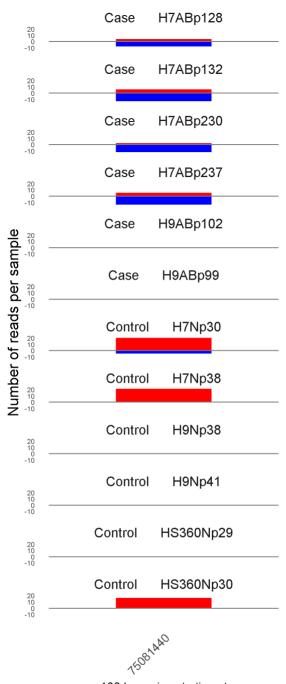






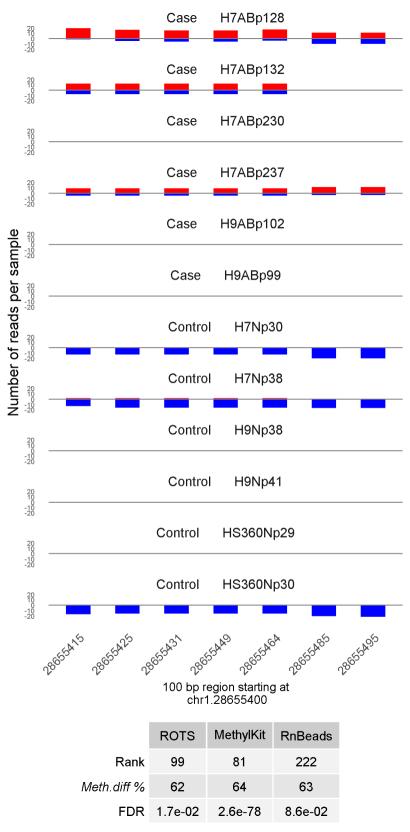


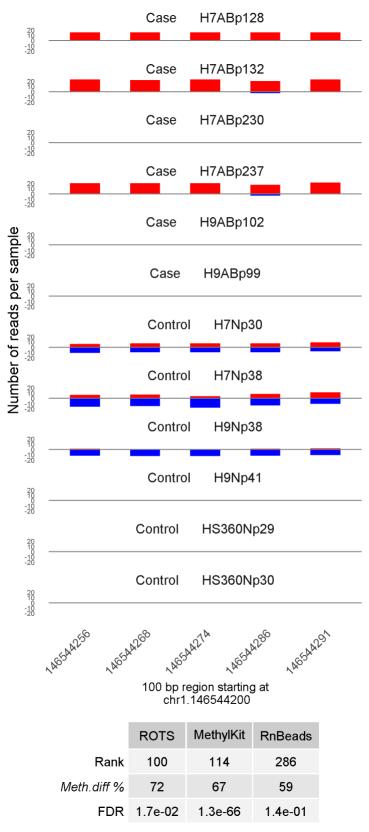


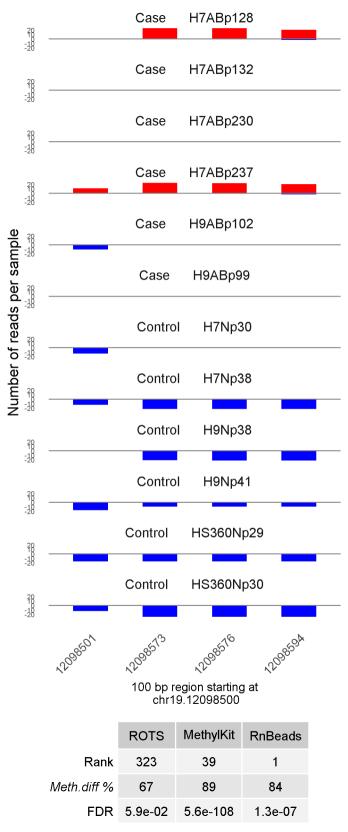


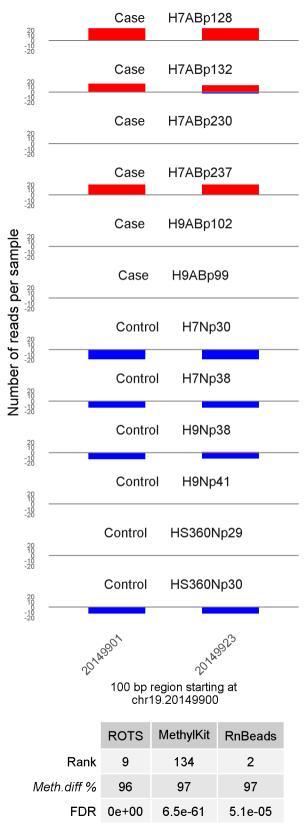
100 bp region starting at chr15.75081400

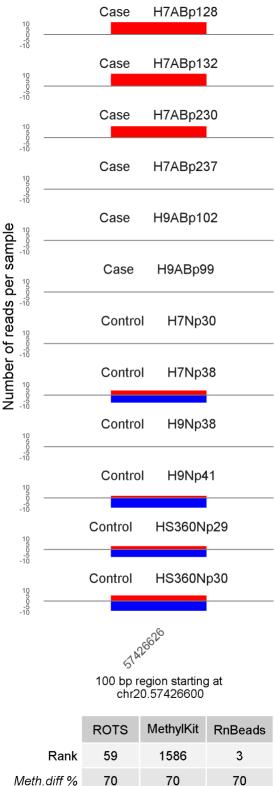
	ROTS	MethylKit	RnBeads
Rank	98	1230	805
Meth.diff %	-67	-65	-67
FDR	1.7e-02	3.5e-13	4.6e-01





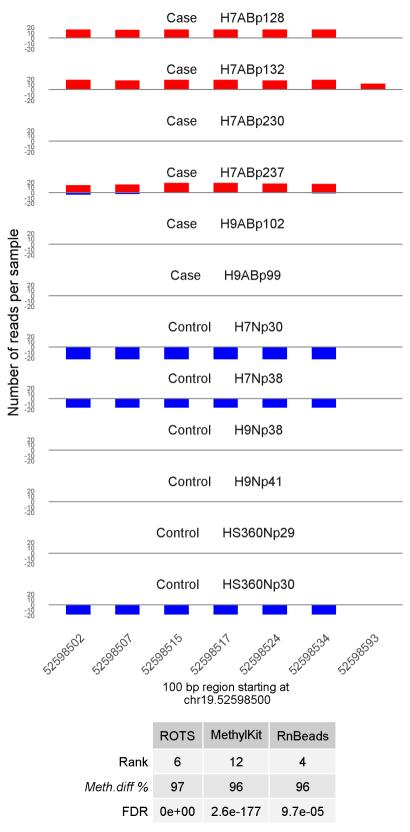


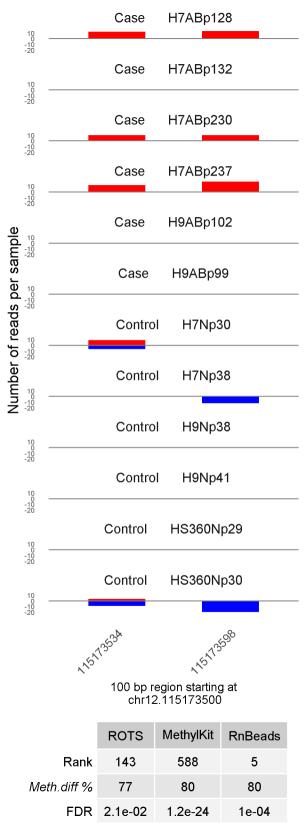


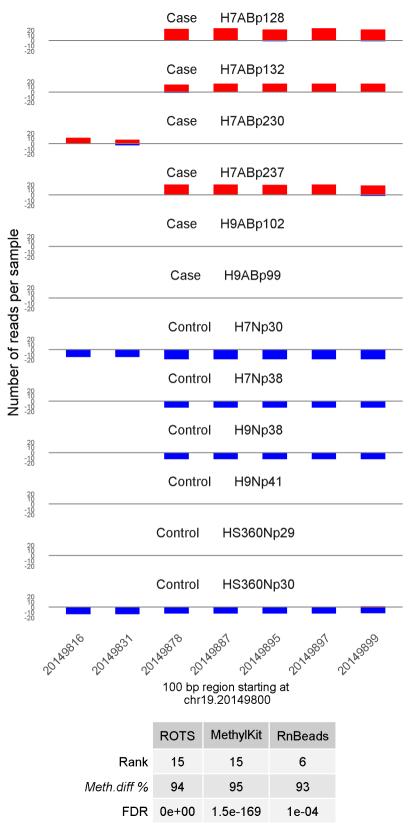


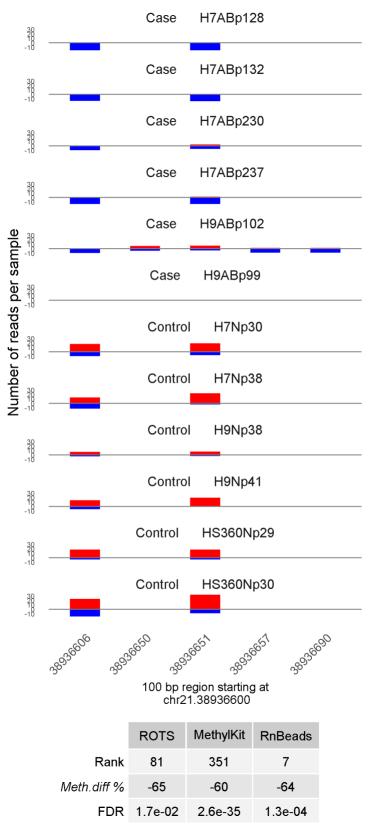
FDR 1.3e-02 1.9e-10

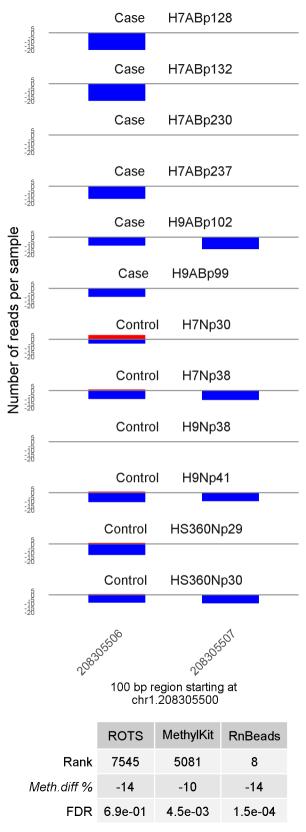
6.3e-05

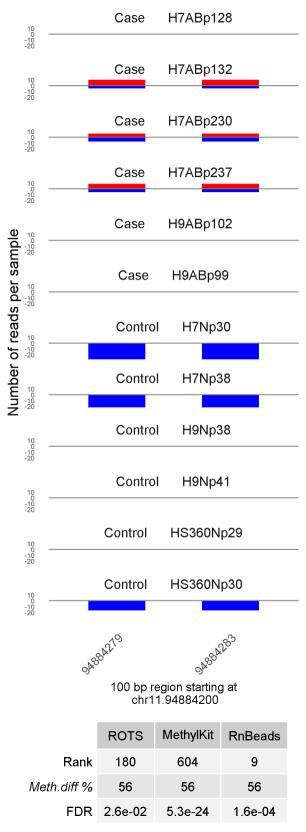


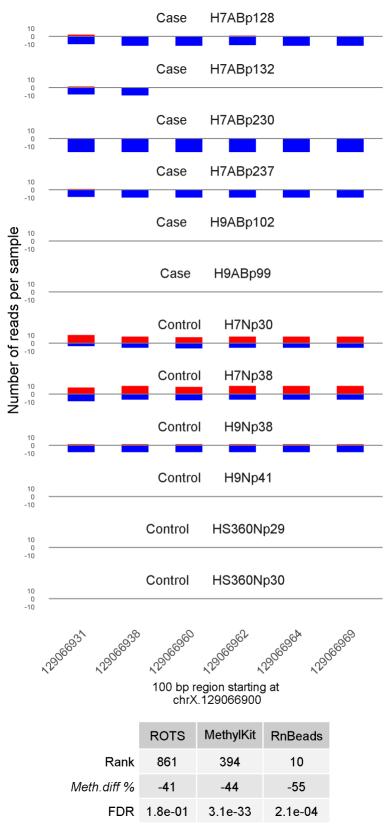


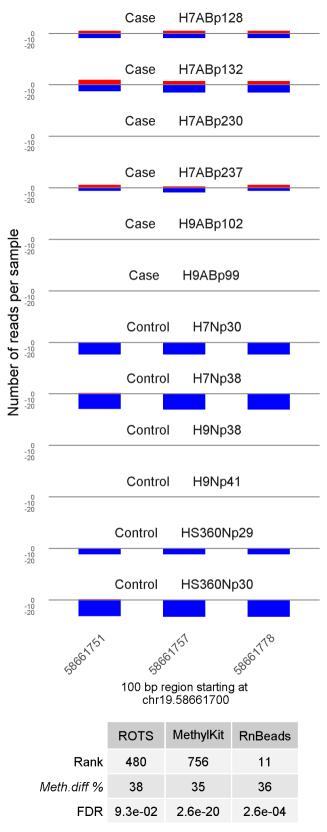


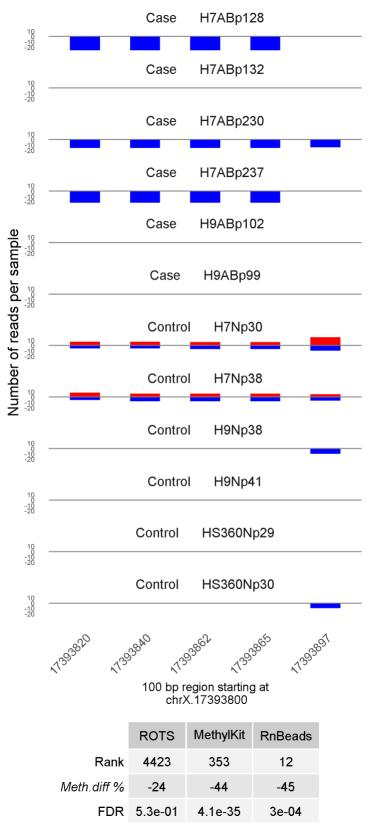


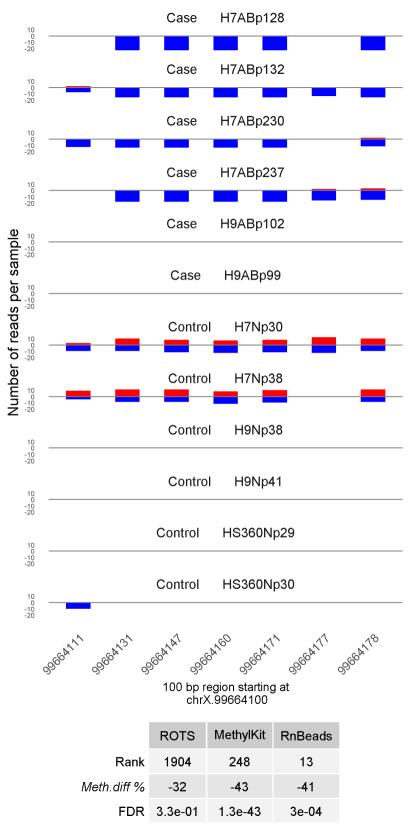


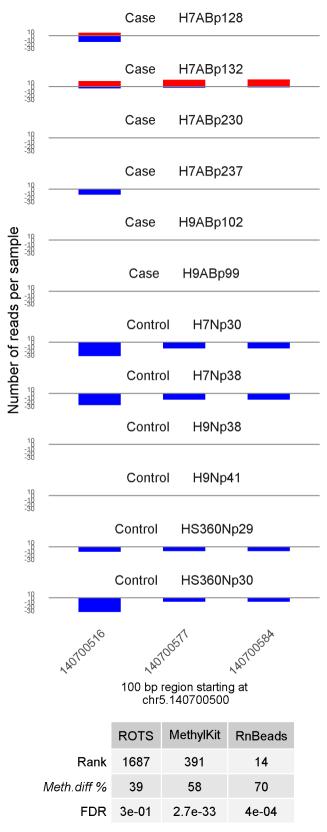


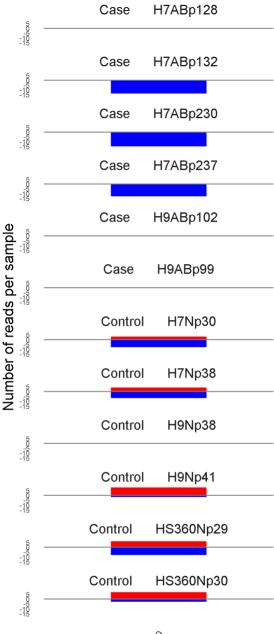








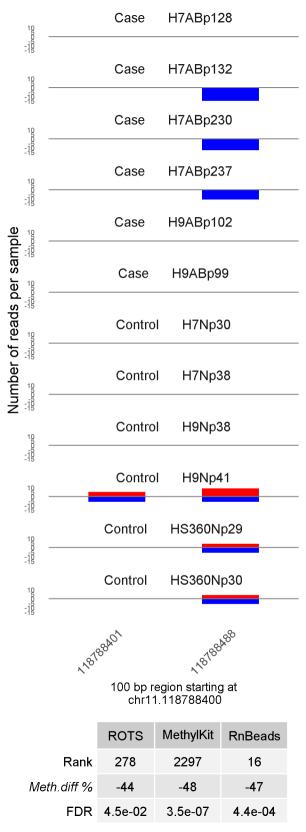


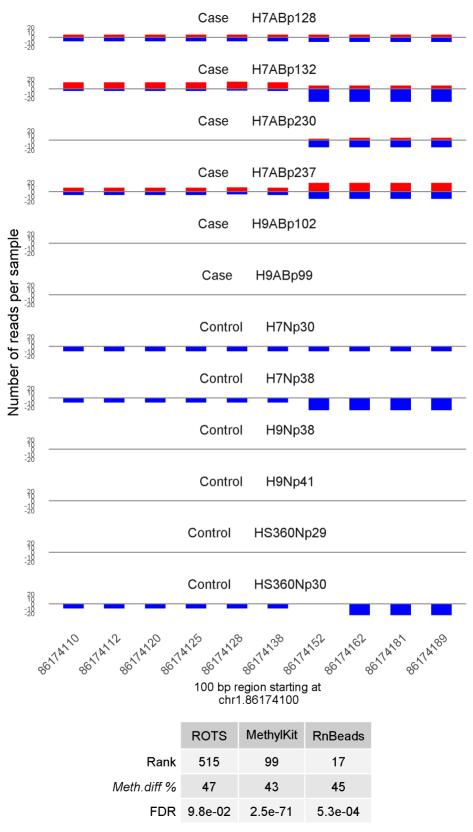


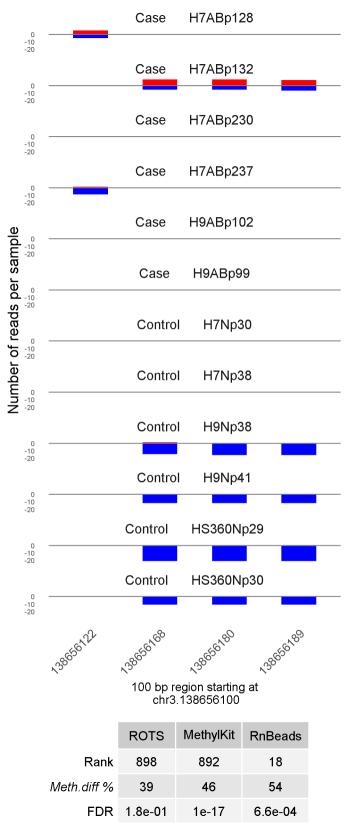


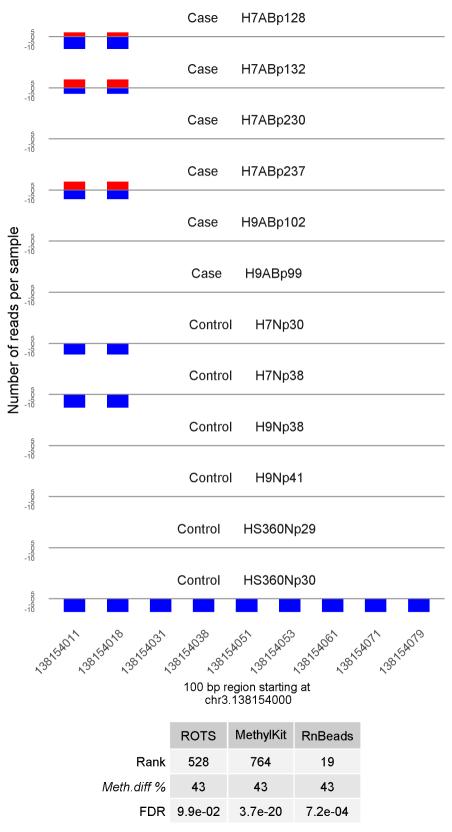
100 bp region starting at chr1.2564700

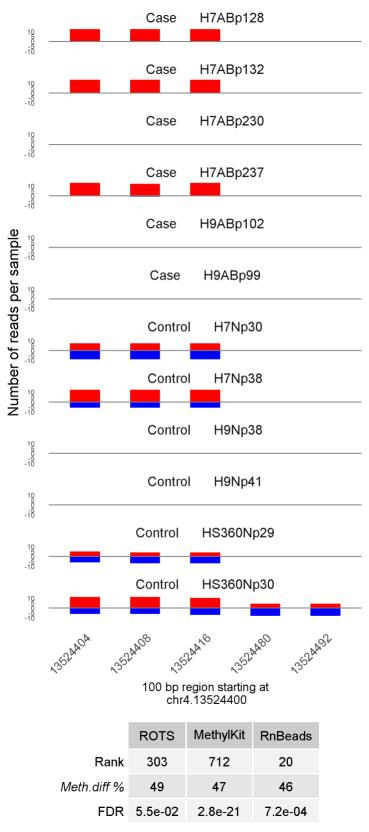
	ROTS	MethylKit	RnBeads
Rank	411	1945	15
Meth.diff %	-51	-50	-51
FDR	7.5e-02	1.9e-08	4.1e-04

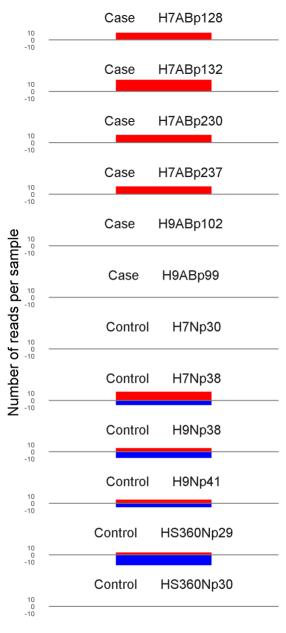








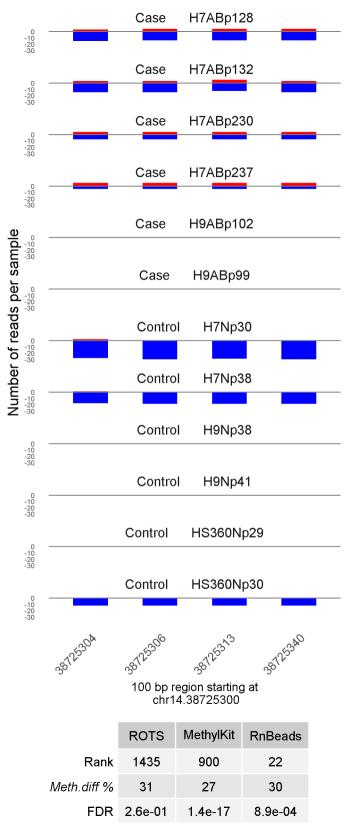


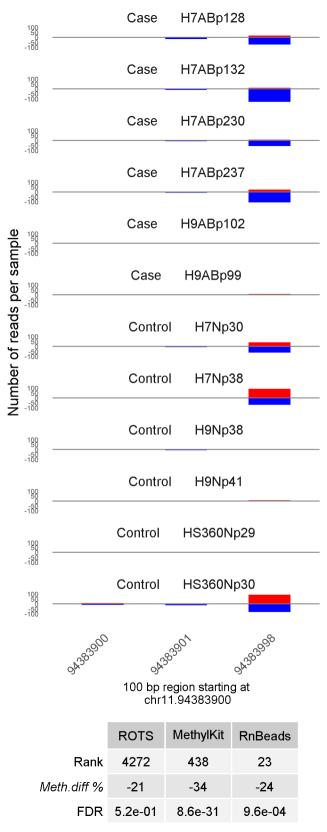


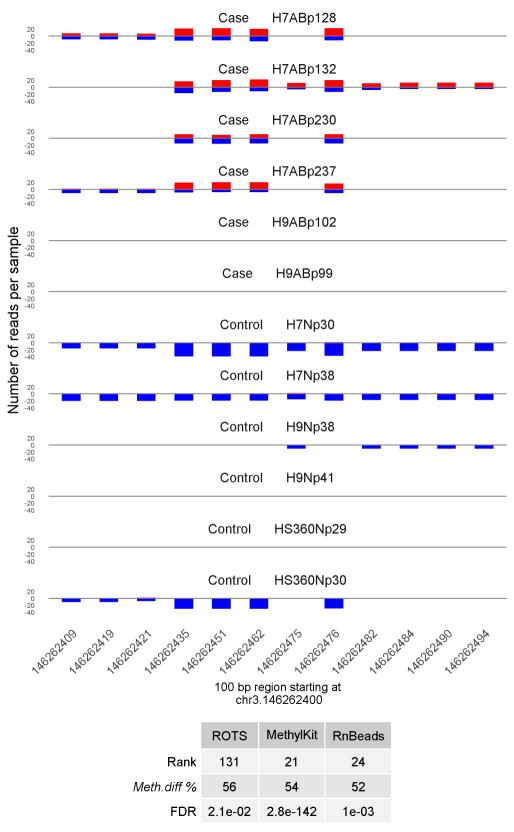


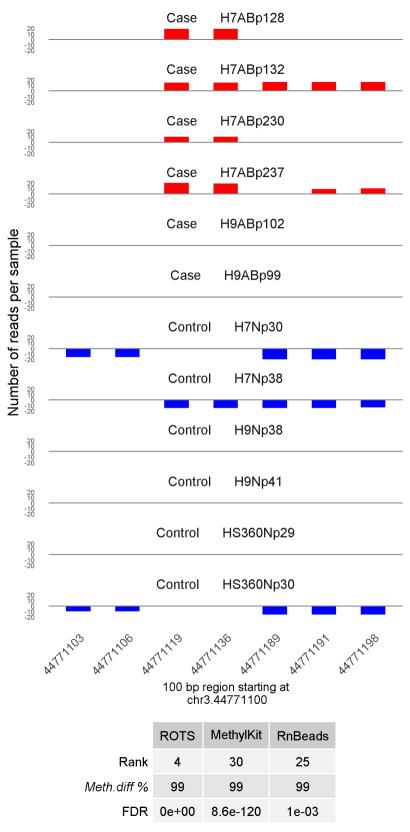
100 bp region starting at chr4.13525300

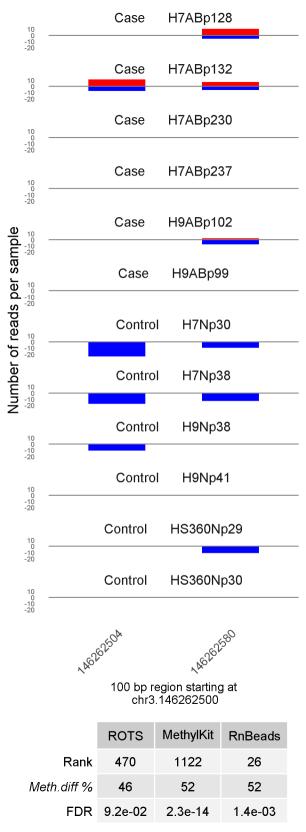
	ROTS	MethylKit	RnBeads
Rank	187	1357	21
Meth.diff %	60	60	58
FDR	2.6e-02	4.2e-12	7.7e-04

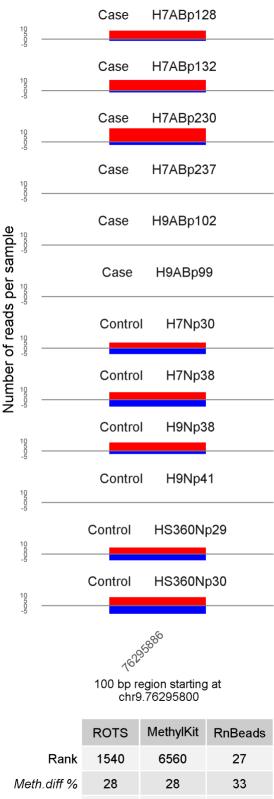






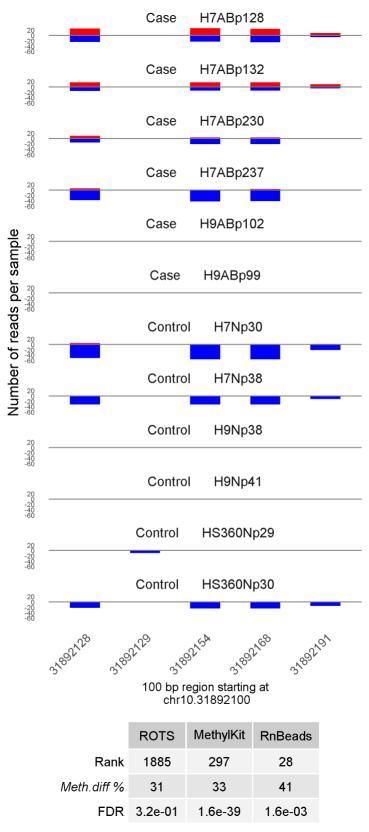


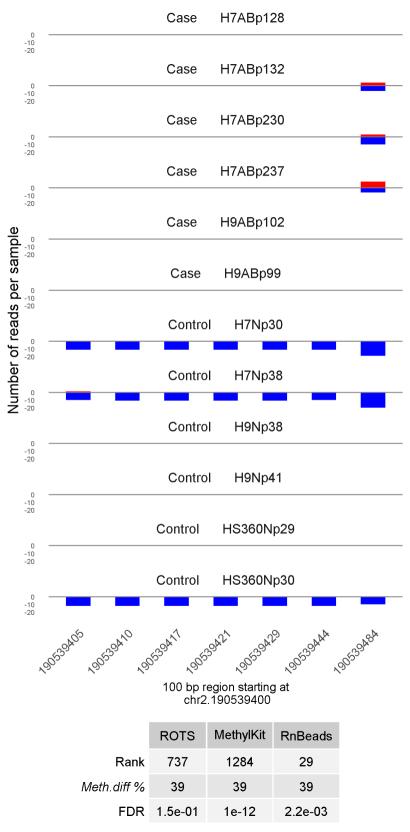


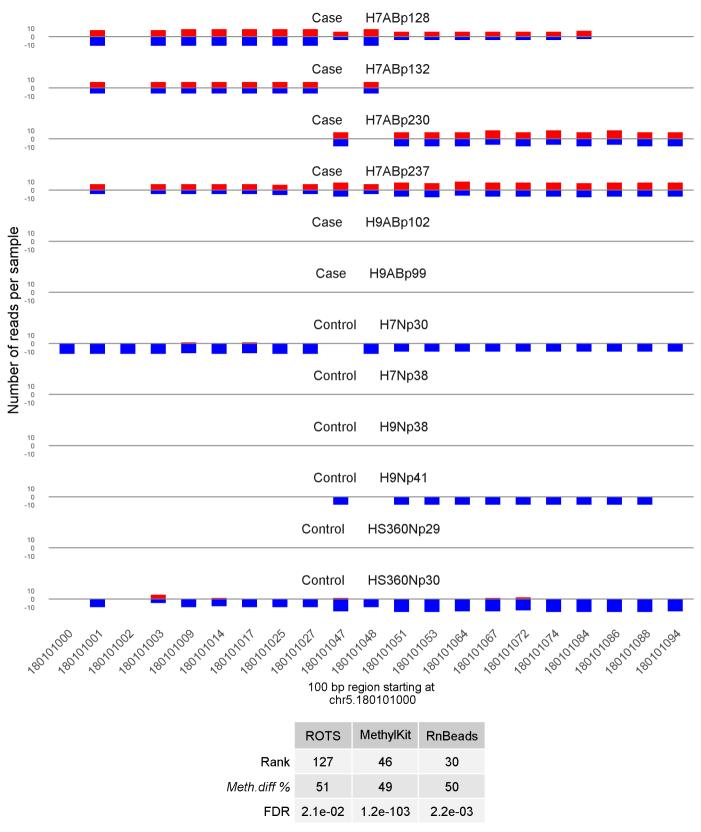


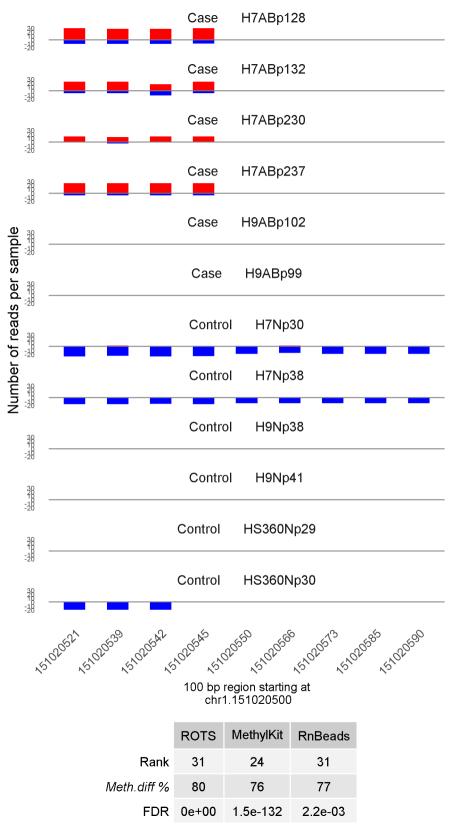
FDR 2.8e-01 2.6e-02

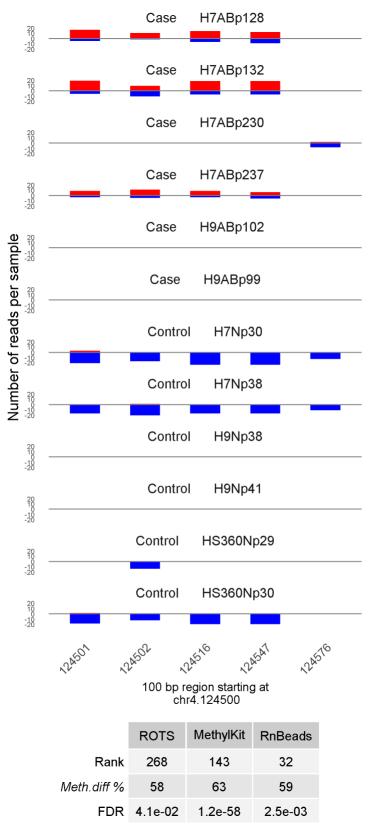
02 1.6e-03

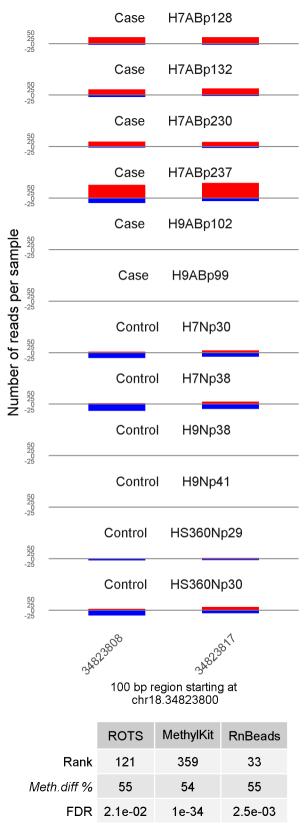


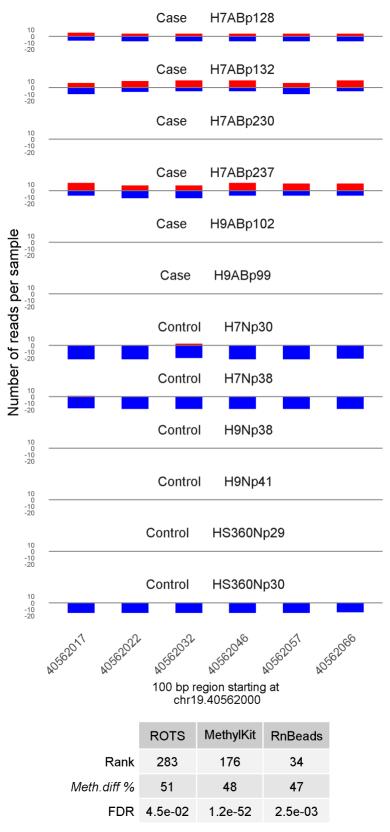


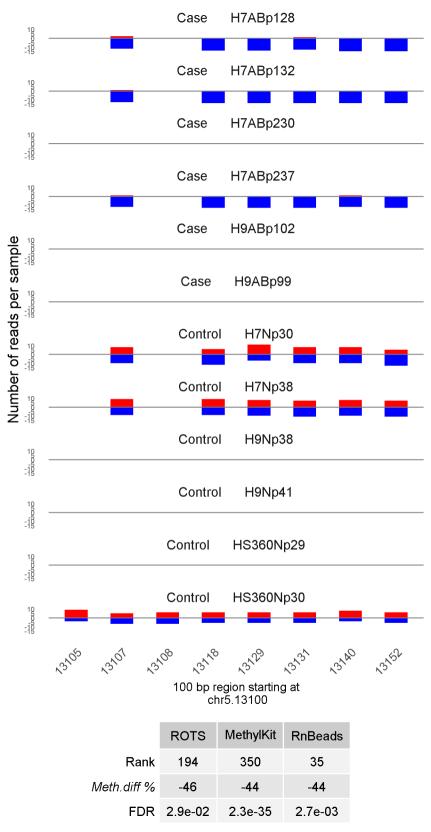


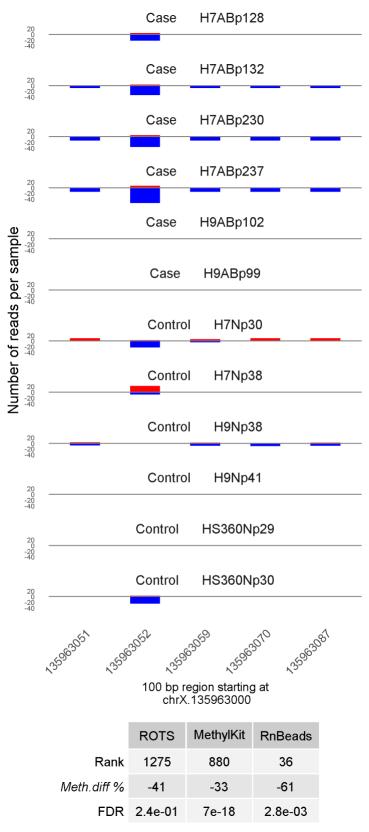


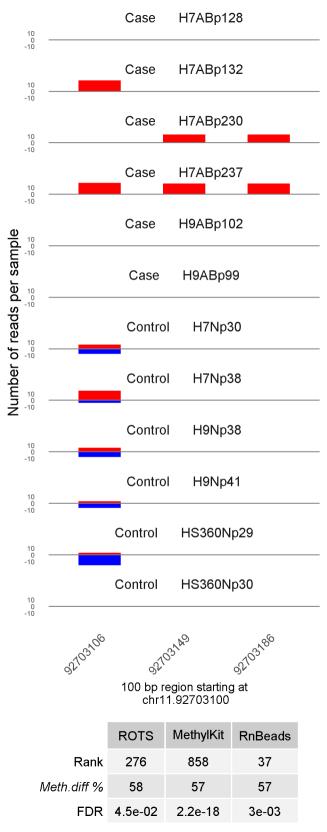




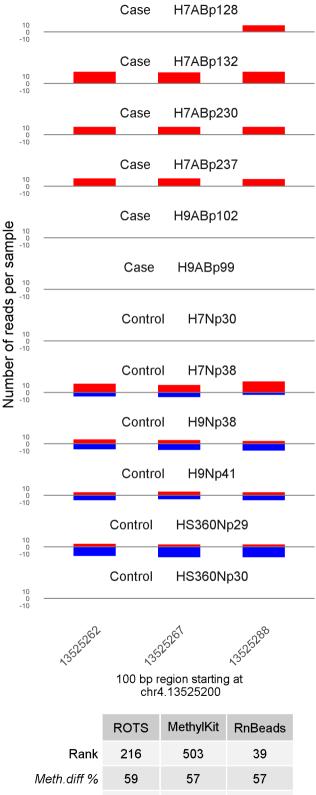




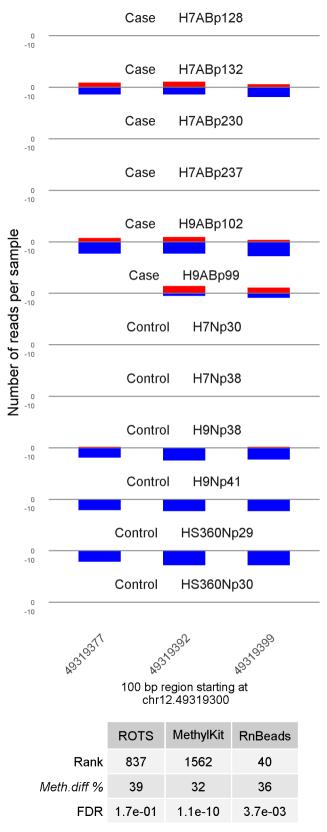


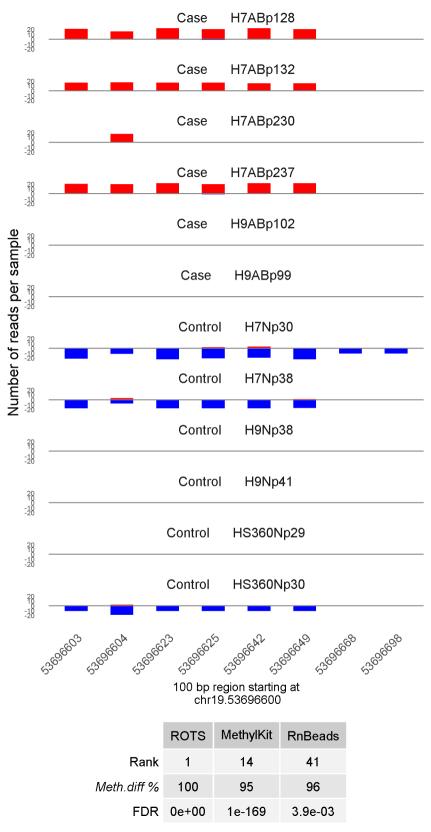


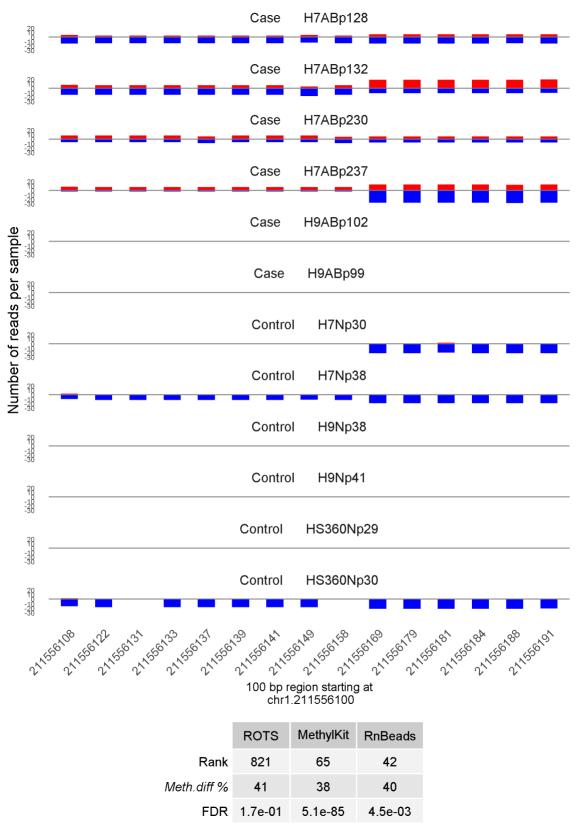
	10		Case	H7ABp12	28	
	10 -10 -20 10 -10 -10 -20		Case	H7ABp13	32	
Number of reads per sample	-10 -20 10 -10 -20		Case	H7ABp23	30	
	-20 10 -10 -20		Case	H7ABp23	37	
			Case	H9ABp10)2	
	-10 -20		Case	Н9АВр9	9	
	-10 -20		Contro	I H7Np3	60	
	0 -10 -20		Contro	I H7Np3	8	
			Contro	I H9Np3	8	
	10 -10 -20		Contro	I H9Np4	.1	
	10 0 -10 -20		Control	HS360N	p29	
	10 -10 -20		Control	HS360N	p30	
	10 0 -10 -20		.100		6 2	
		63 ⁵⁶¹⁰⁰ 63 ⁵⁶¹⁷⁵⁵ 100 bp region starting at chr19.53661700				
Rank <i>Meth.diff %</i> FDR		ROTS	MethylKit	RnBeads		
		547	1775	38		
		50	54	53		
		1.1e-01	2.3e-09	3e-03		

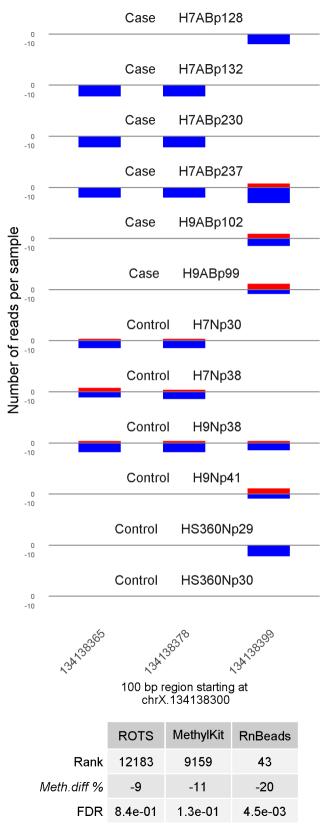


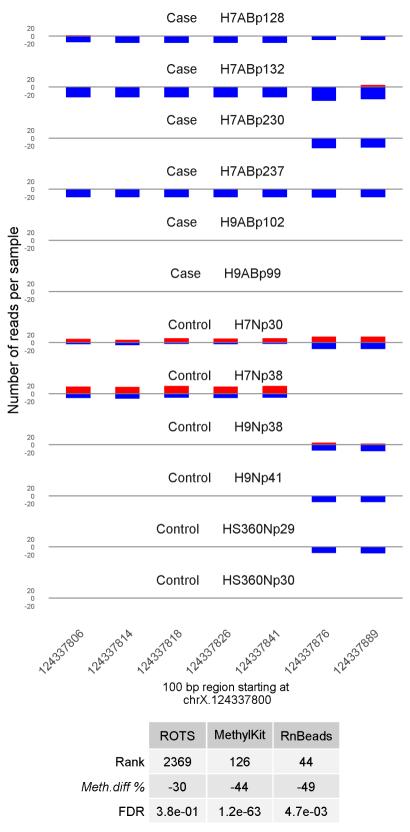
FDR 3.2e-02 8.1e-28 3.3e-03

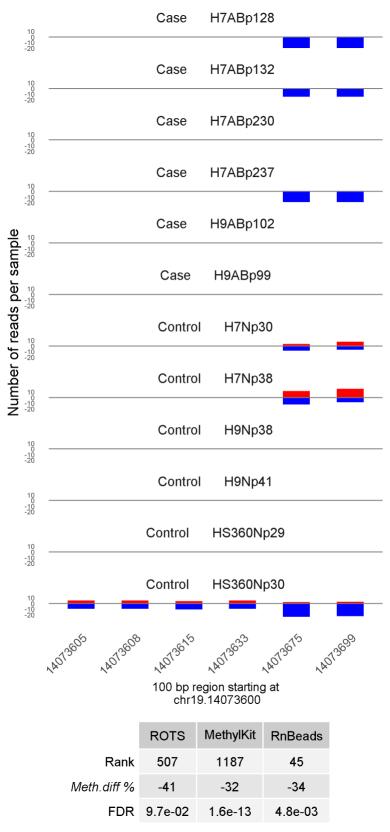


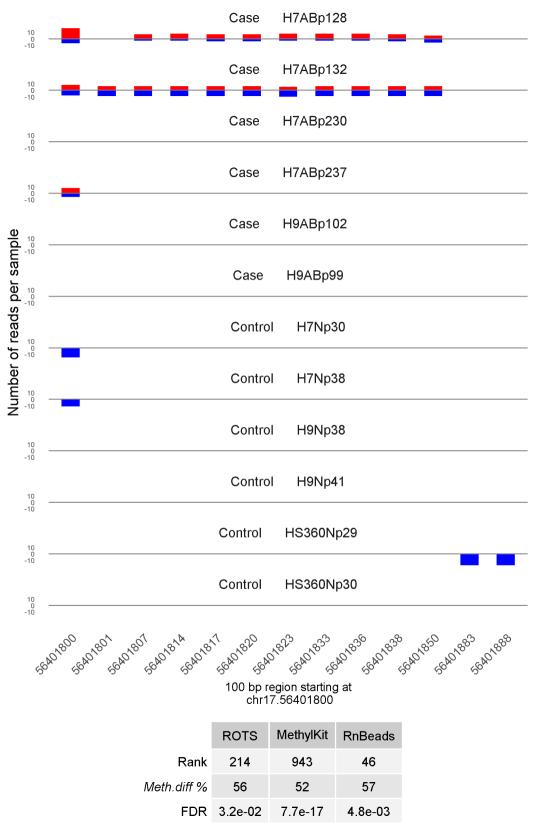


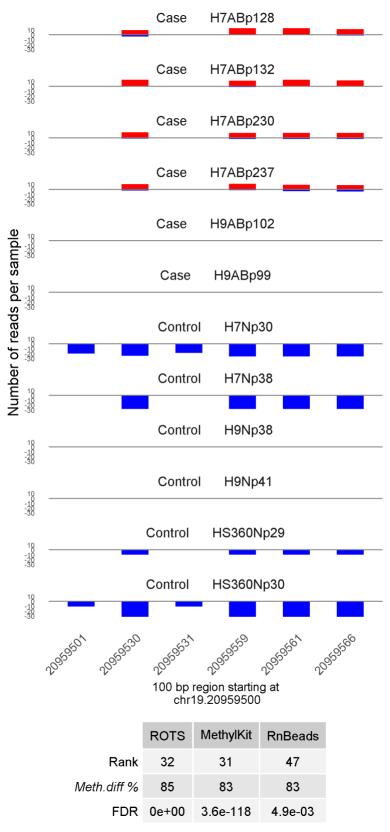


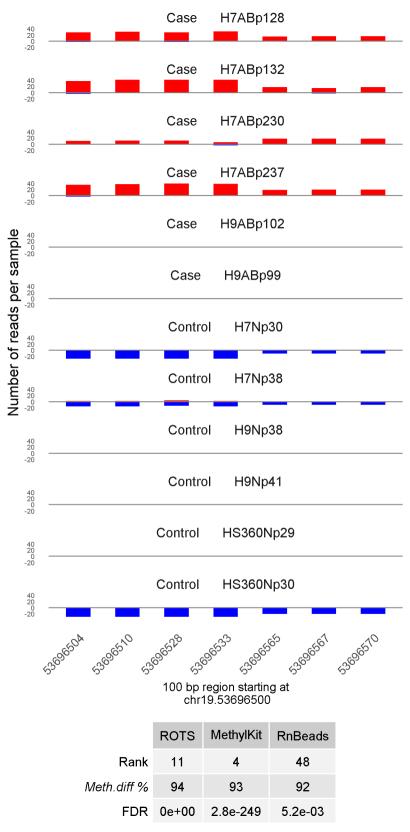


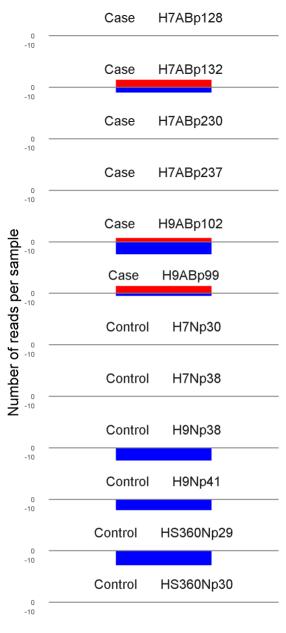








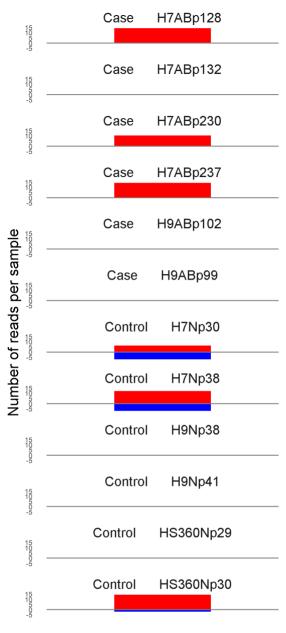






100 bp region starting at chr12.49319400

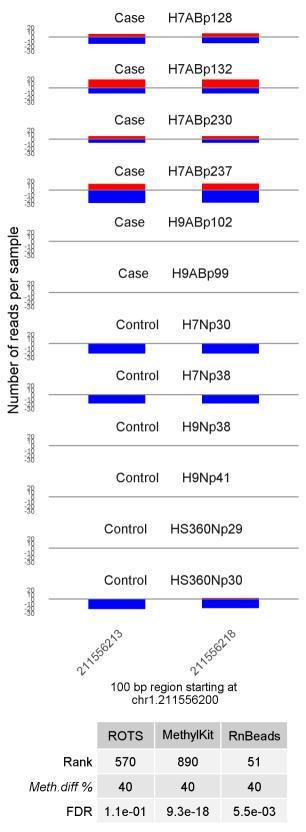
	ROTS	MethylKit	RnBeads
Rank	481	2316	49
Meth.diff %	51	47	51
FDR	9.3e-02	3.8e-07	5.3e-03

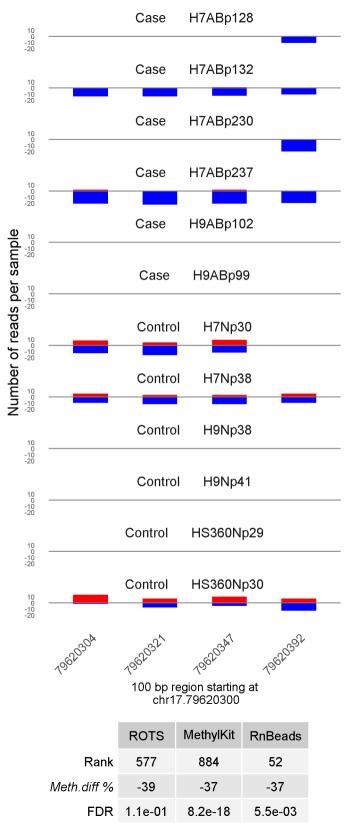


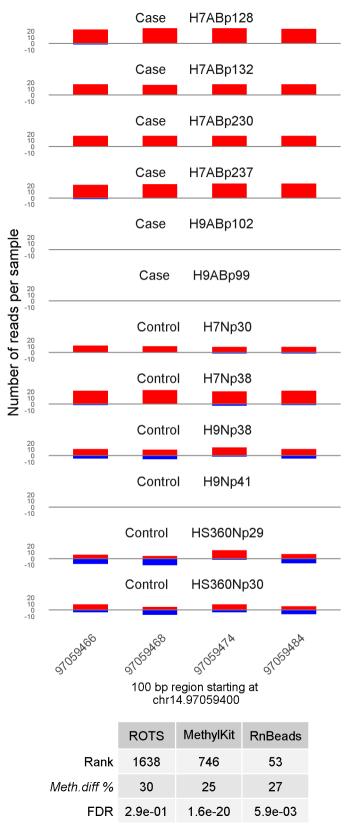


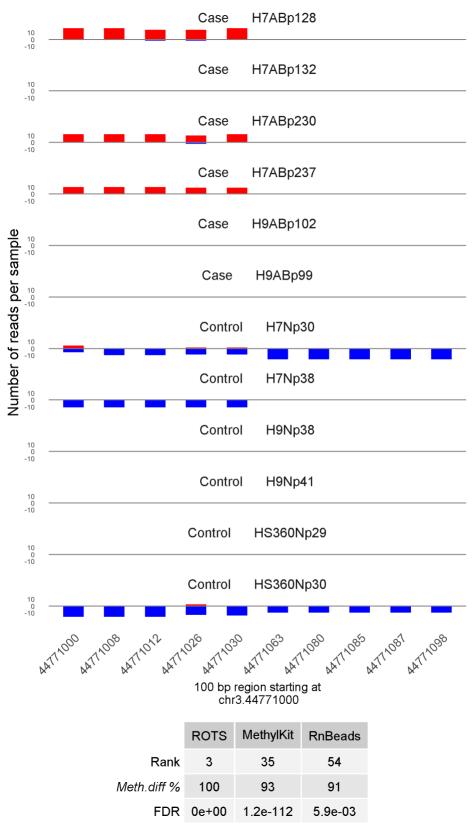
100 bp region starting at chr11.75491900

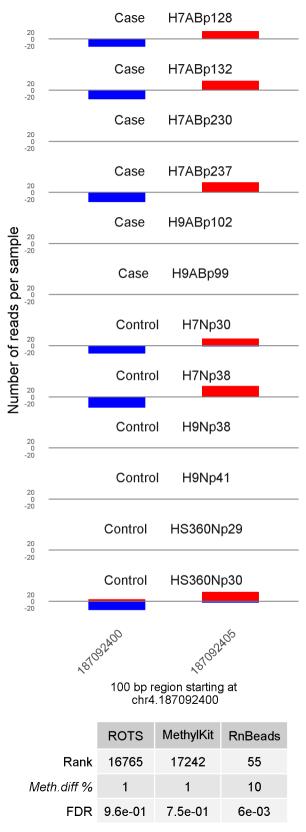
	ROTS	MethylKit	RnBeads
Rank	1321	3420	50
Meth.diff %	34	33	34
FDR	2.5e-01	1e-04	5.4e-03

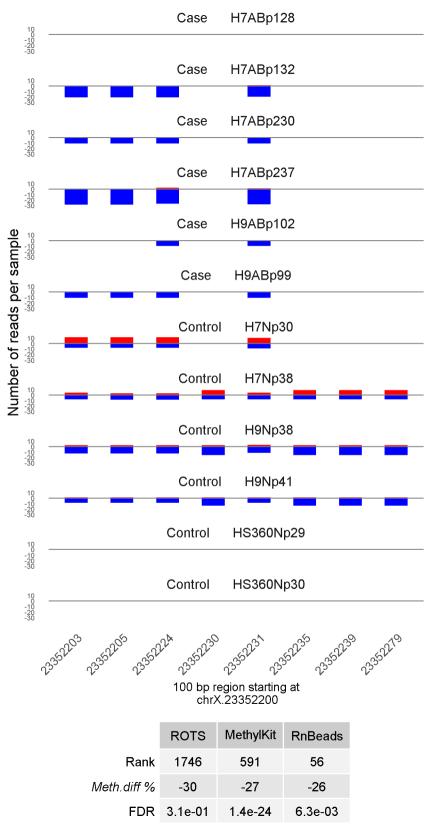


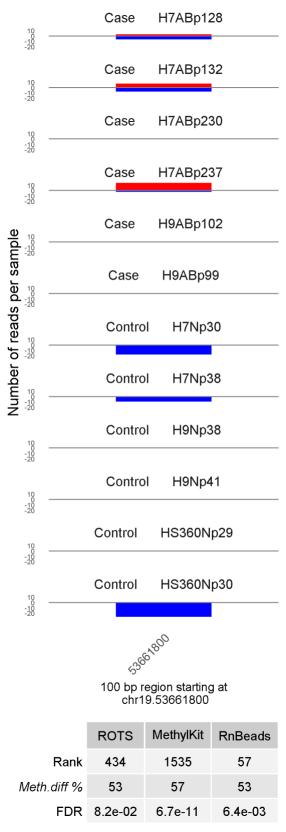


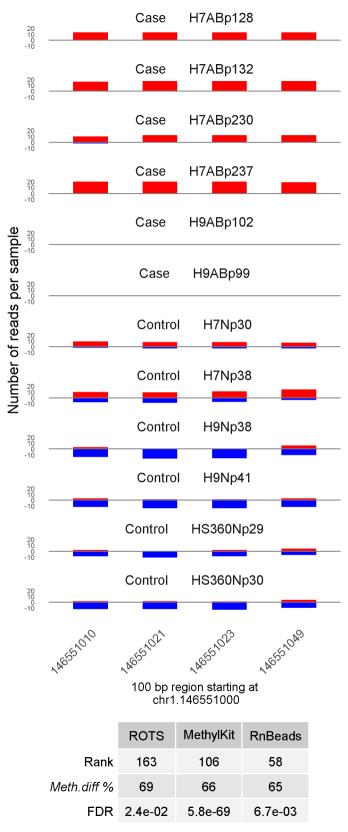


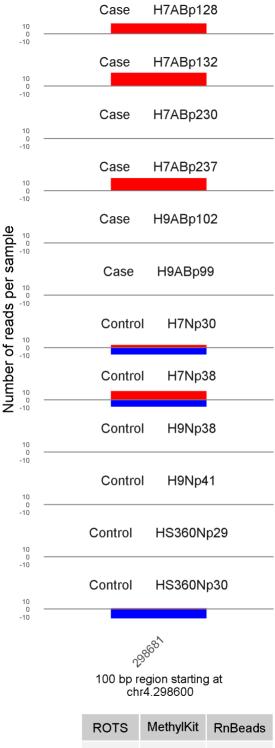




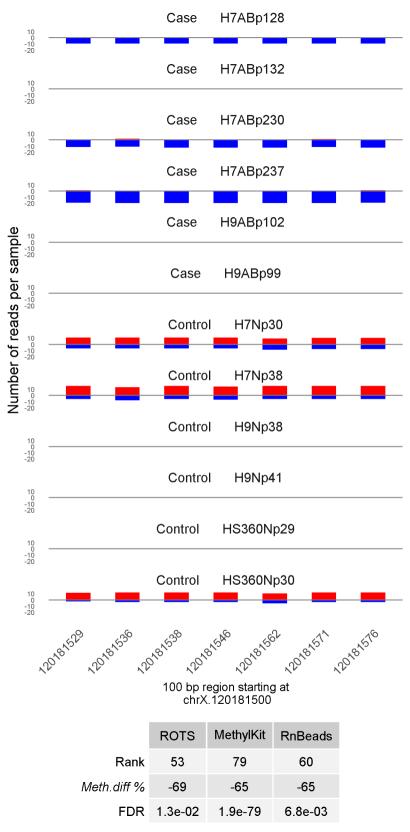


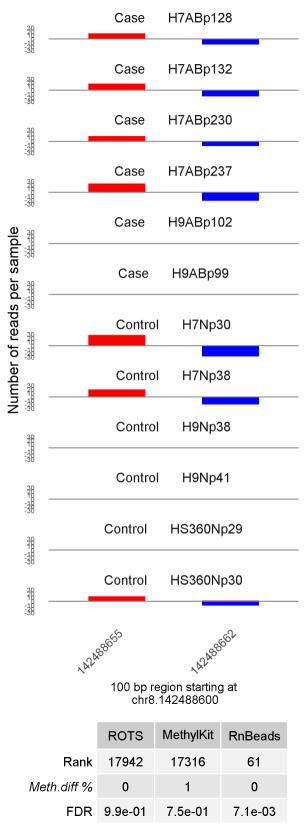


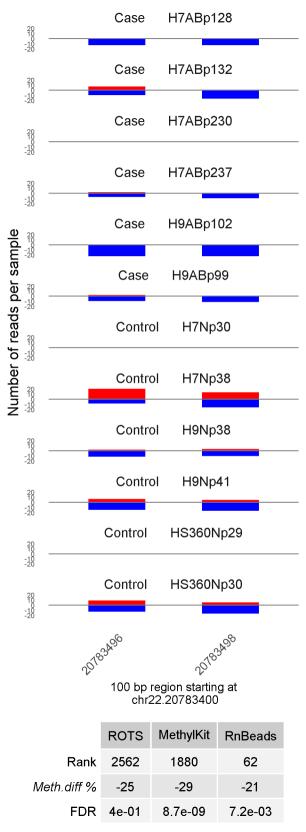


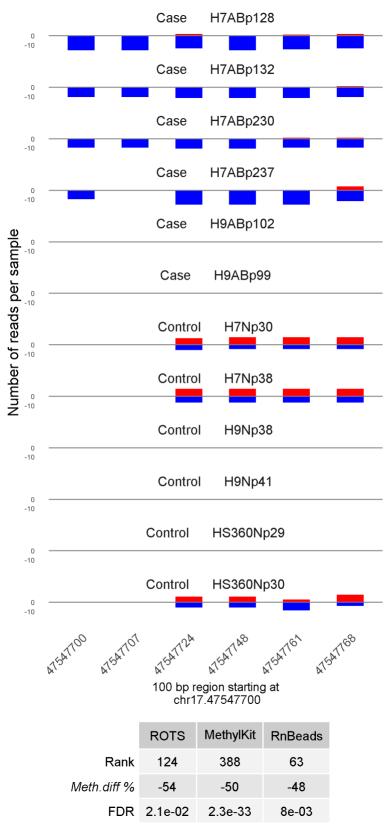


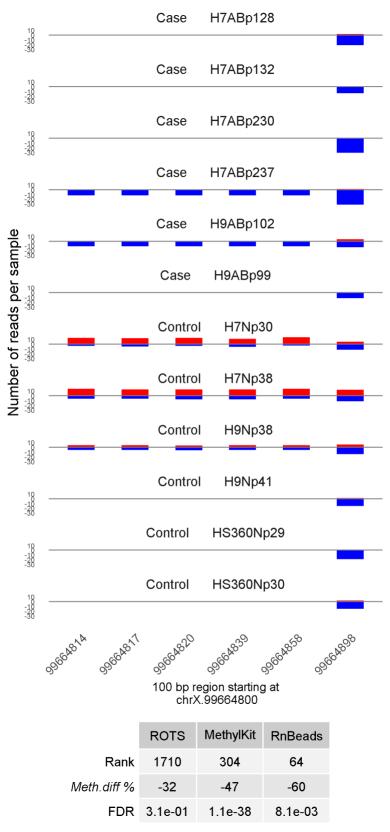
Rank	148	1327	59
Meth.diff %	71	67	71
FDR	2.4e-02	2.3e-12	6.7e-03

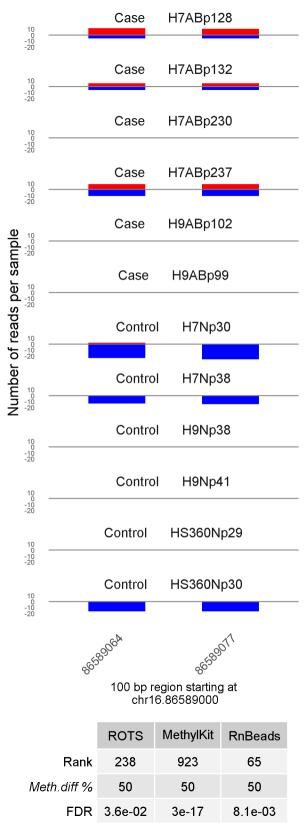


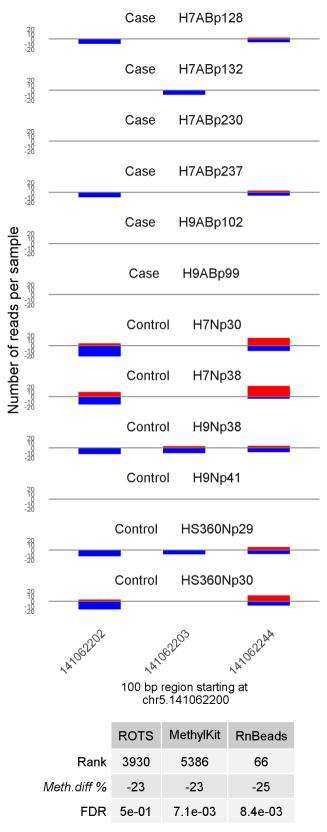


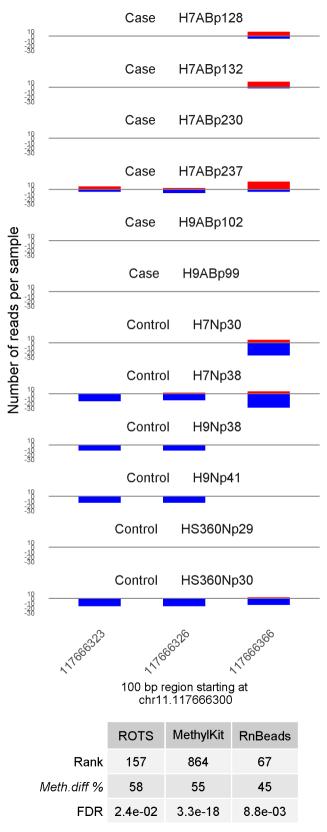


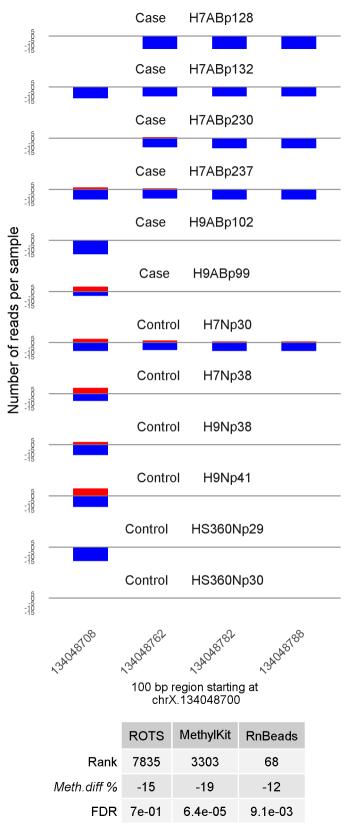


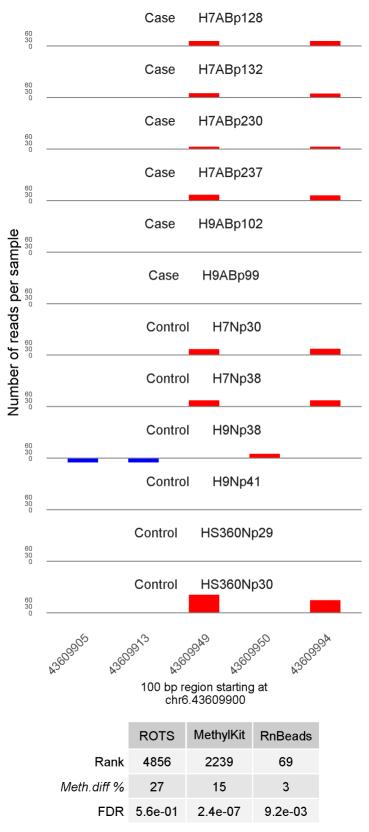


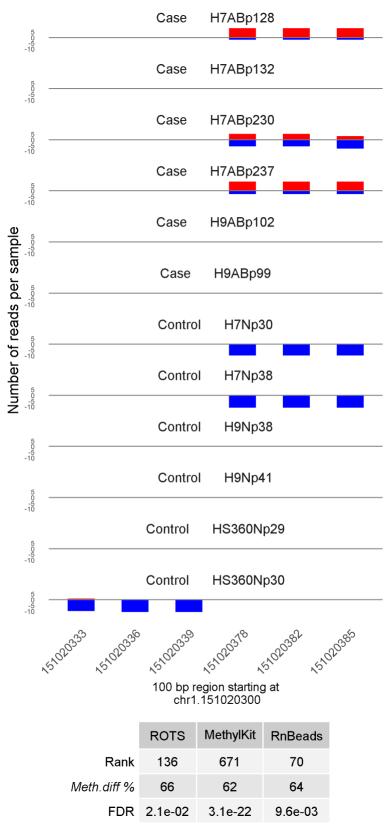


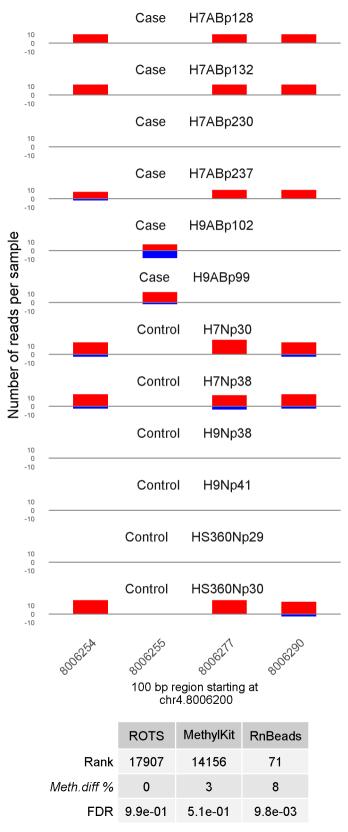


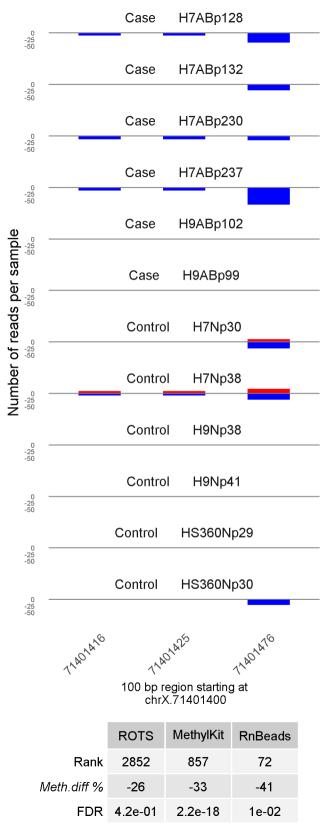


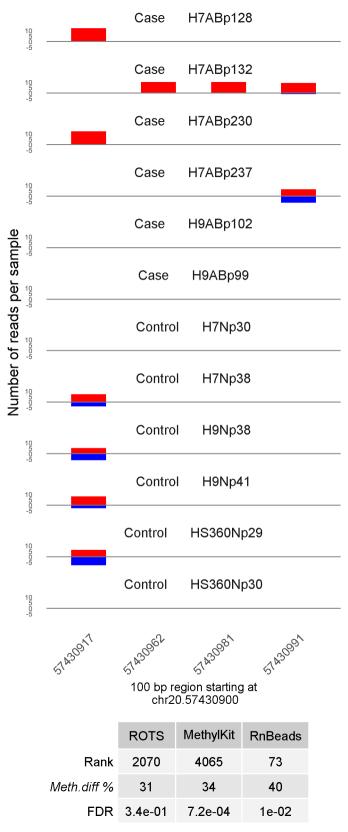


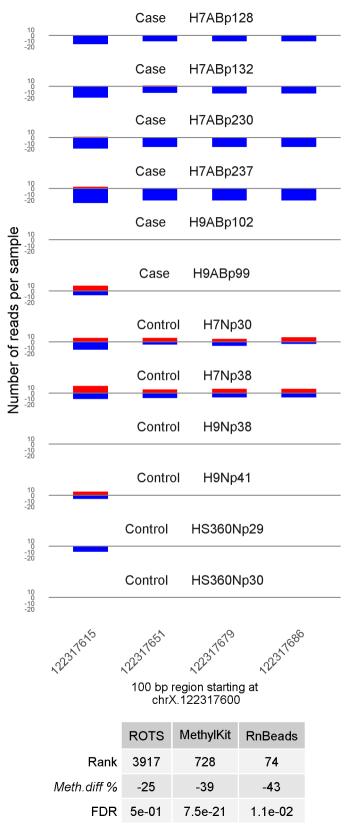


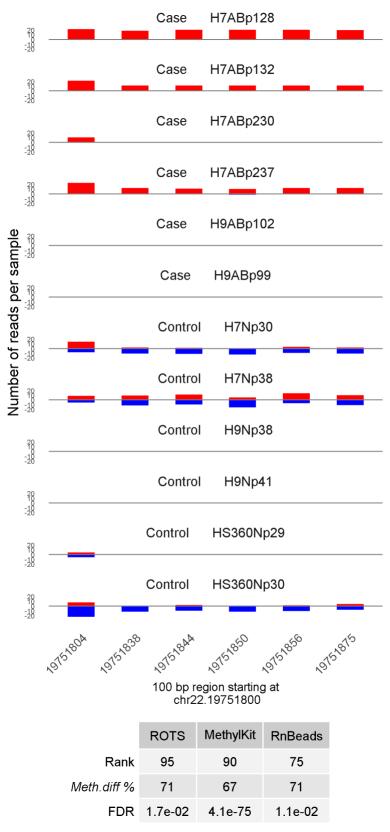


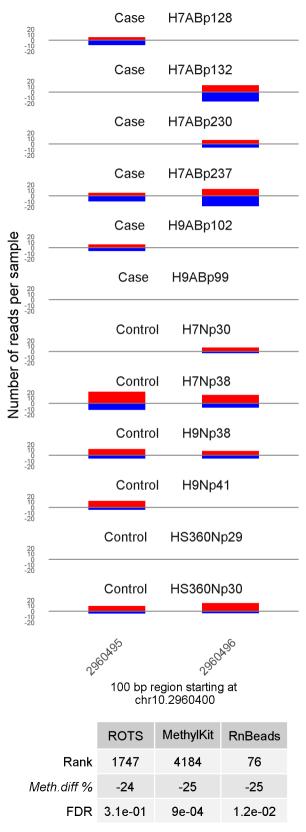




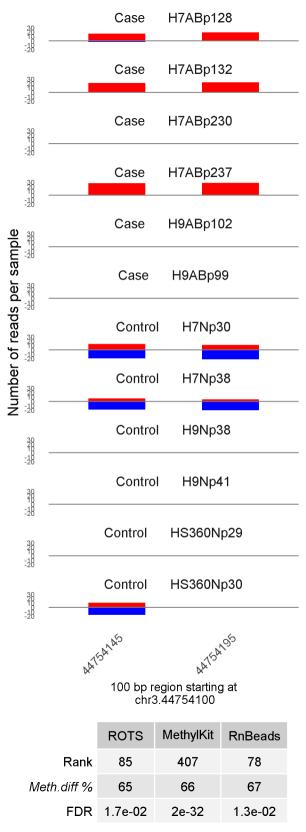


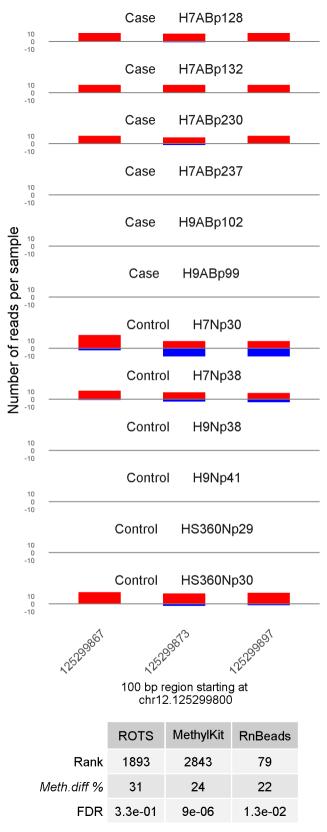


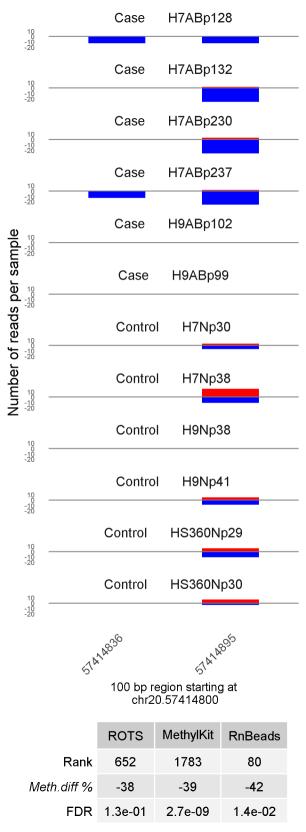


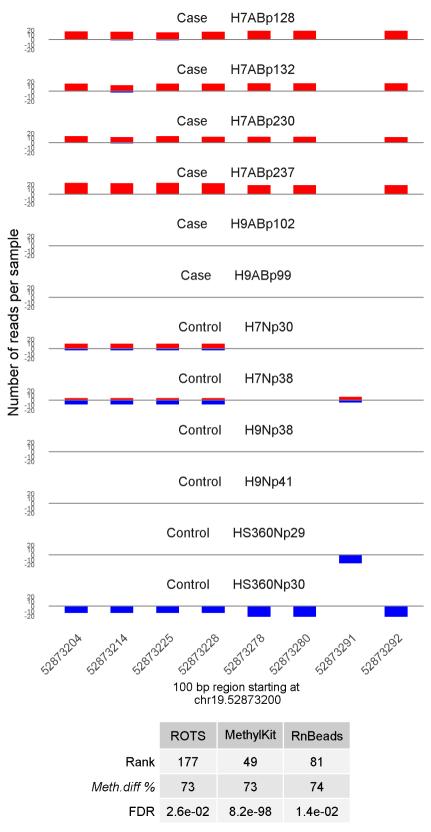


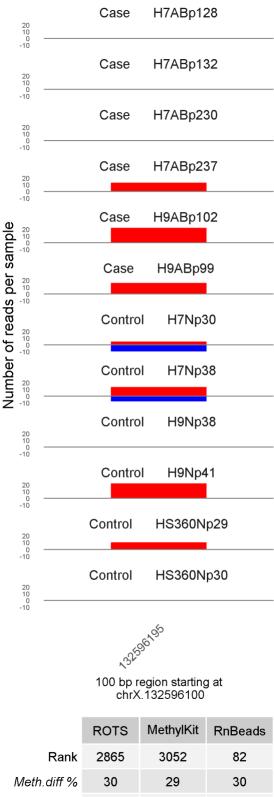
	25		Case	H7ABp1	28		
	25 -25 -50 -75						
	25		Case	H7ABp1	32		
	25 -25 -50 -75						
	25		Case	H7ABp2	30		
	25 -25 -50 -75						
	25		Case	H7ABp2	37		
	25 -25 -50 -75						
e	25 0		Case	H9ABp1	02		
amp	25 -25 -50 -75						
Number of reads per sample	25		Case H9ABp99				
ds p	25 -25 -50 -75						
rea	25		Contro	l H7Np:	30		
er of	25 -25 -50 -75		. .				
ğ	25 -25 -50 -75		Contro	l H7Np3	38		
ž	-50 -75		Cantra				
	25 -25 -50 -75		Contro	l H9Np3	30 		
	-50 -75		Contro	l H9Np4	4.1		
	25 -25 -50 -75				+ 1		
	-75		Control	HS360N	In29		
	25 0-25 -50 -75						
	-75		Control	HS360N	lp30		
	25 -25 -50 -75						
	-75		R		04		
		1. Contraction of the second sec	geble09	\$5959	б` /		
		100 bp region starting at chr17.45952600					
		ROTS	MethylKit	RnBeads			
		Rank	697	1191	77		
Meth.diff % FDR			-34	-36	-36		
			1.4e-01	1.8e-13	1.2e-02		





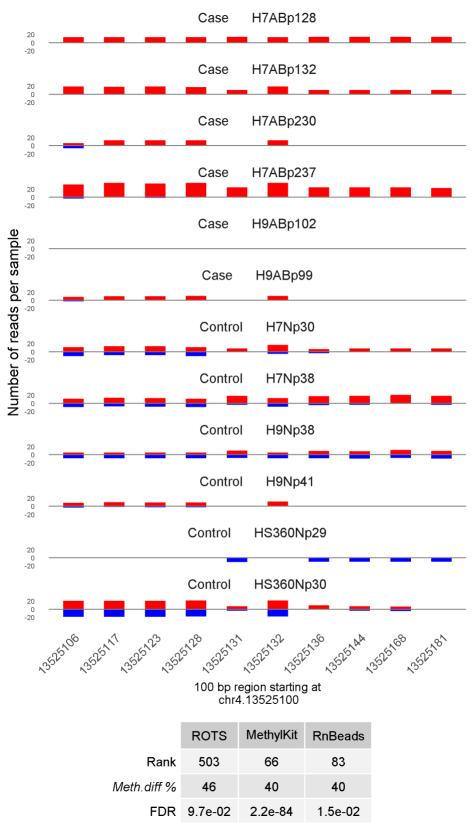


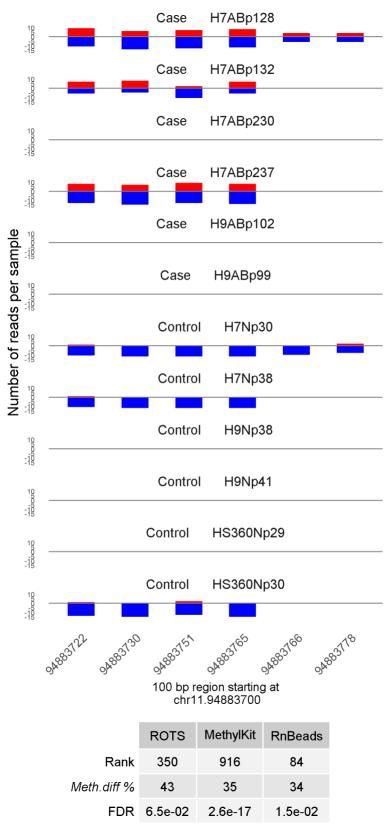


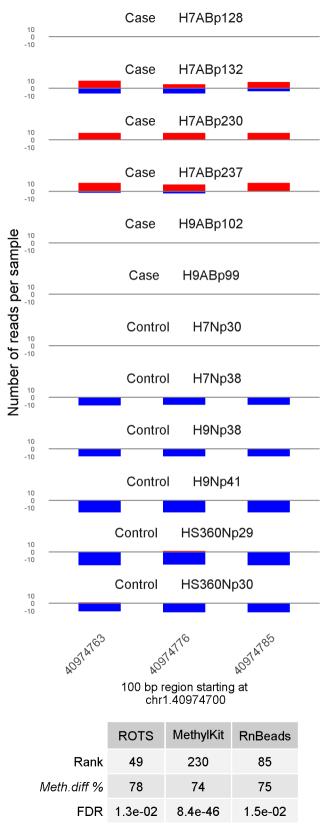


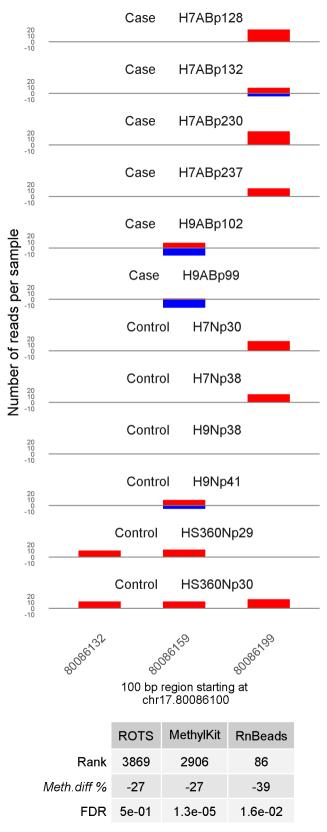
FDR 4.2e-01 2.3

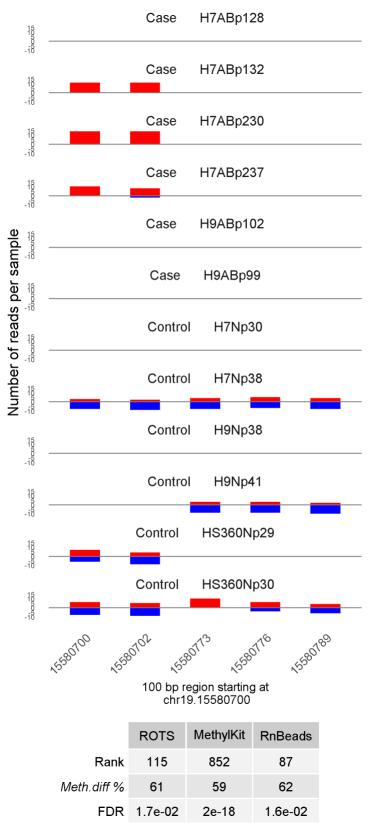
2.3e-05 1.4e-02

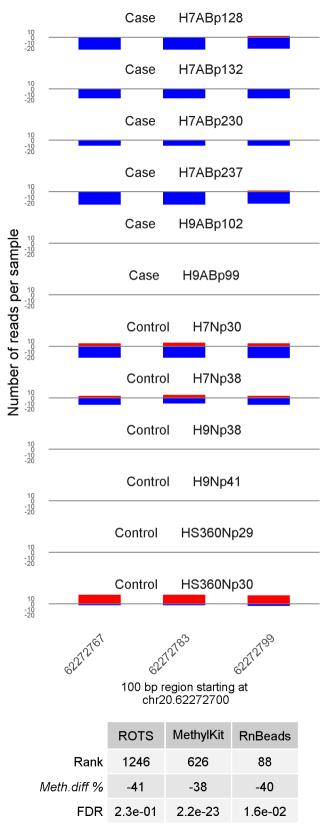


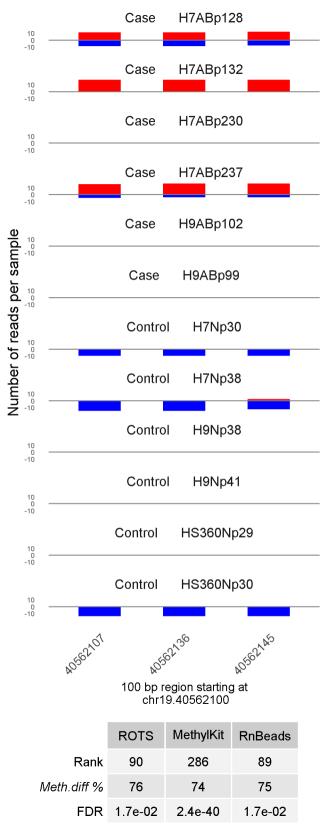


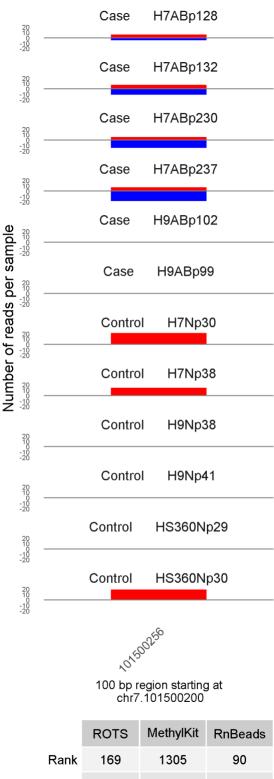








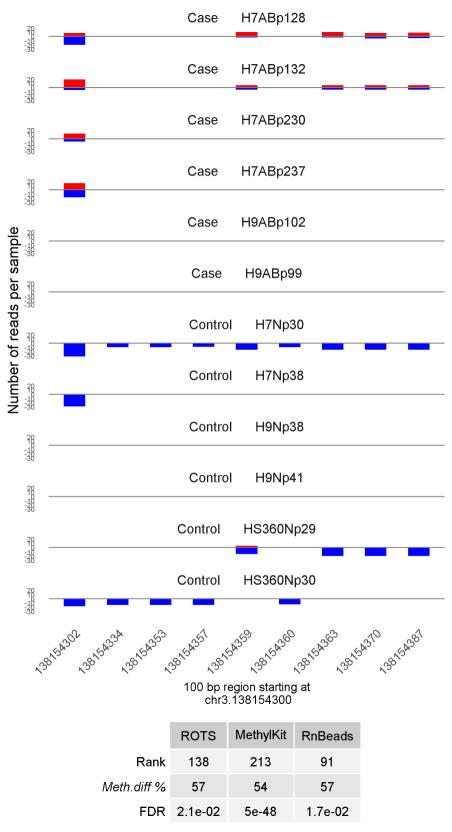


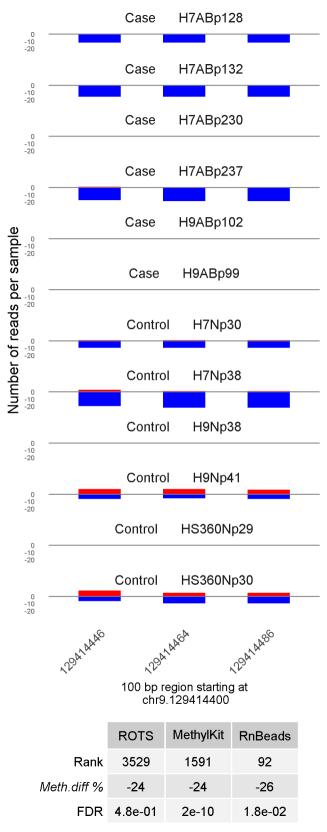


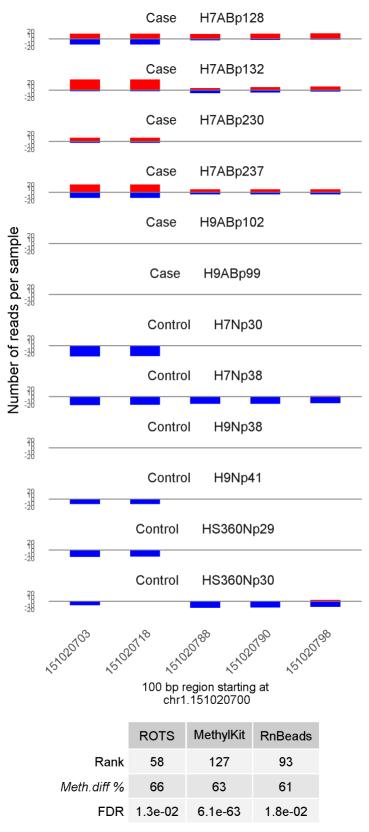
Meth.diff %

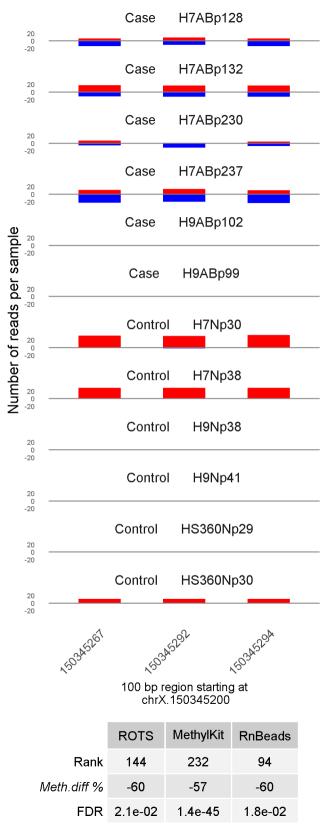
 Iff %
 -58
 -61
 -58

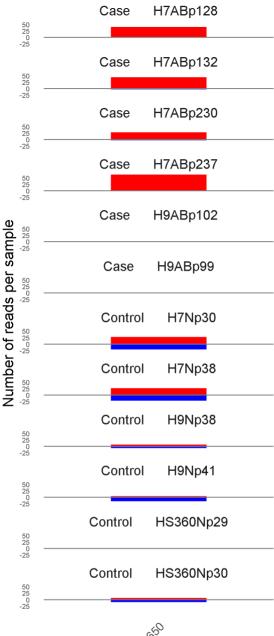
 FDR
 2.6e-02
 1.6e-12
 1.7e-02







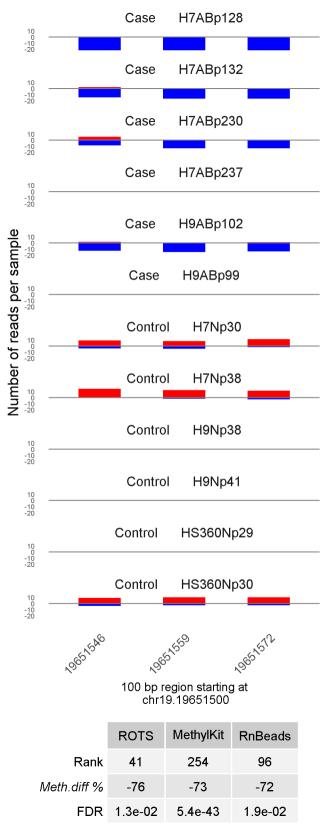


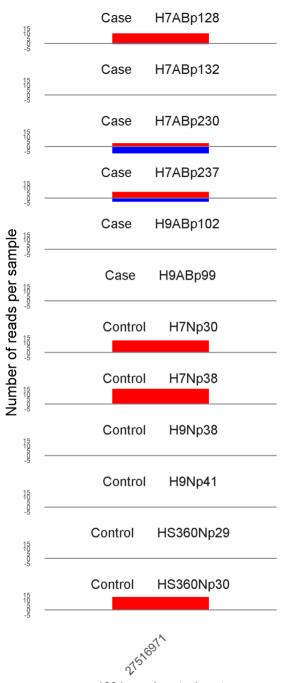




100 bp region starting at chrX.102809600

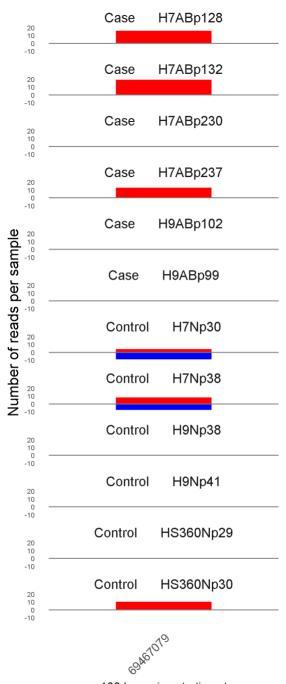
	ROTS	MethylKit	RnBeads
Rank	236	691	95
Meth.diff %	52	48	51
FDR	3.6e-02	1.1e-21	1.8e-02





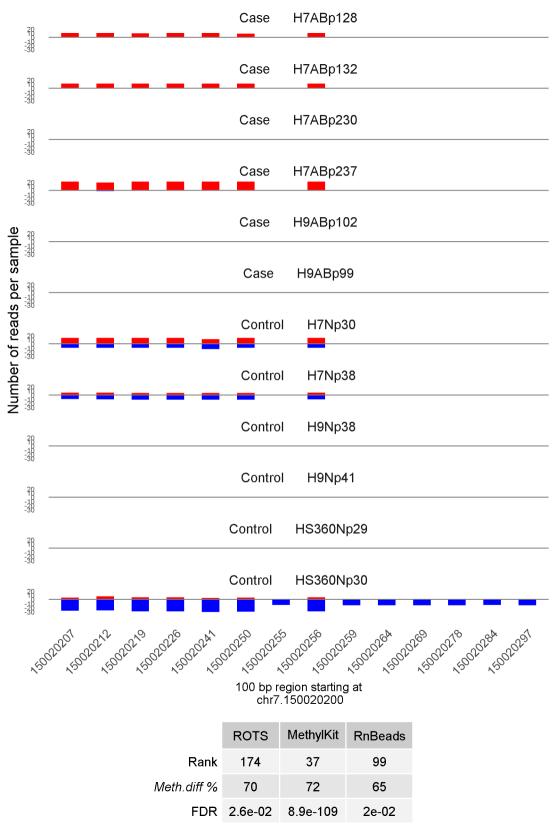
100 bp region starting at chr6.27516900

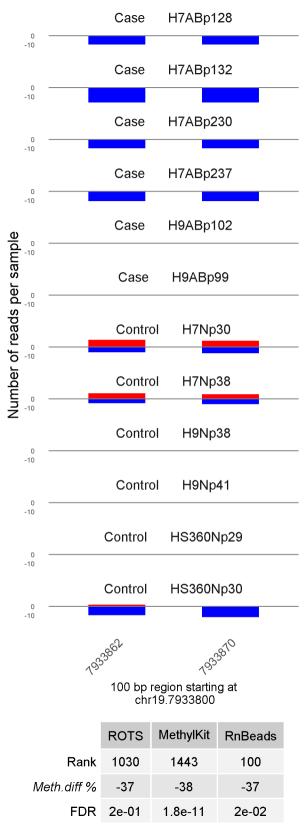
	ROTS	MethylKit	RnBeads
Rank	1291	3221	97
Meth.diff %	-40	-39	-40
FDR	2.4e-01	4.8e-05	1.9e-02

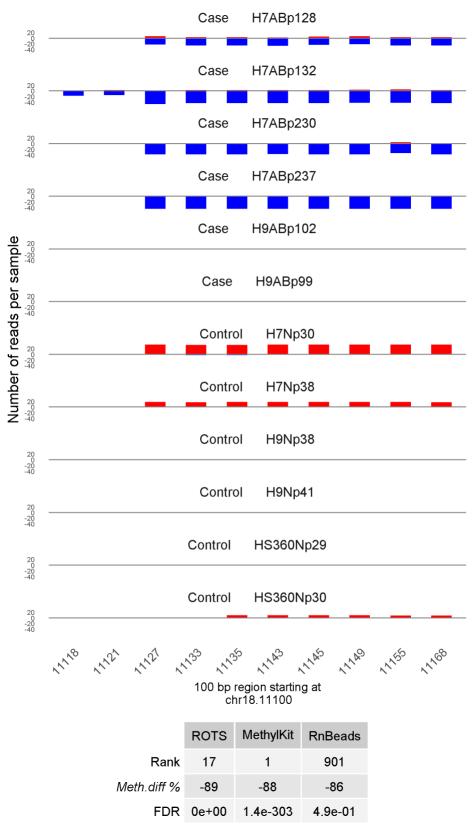


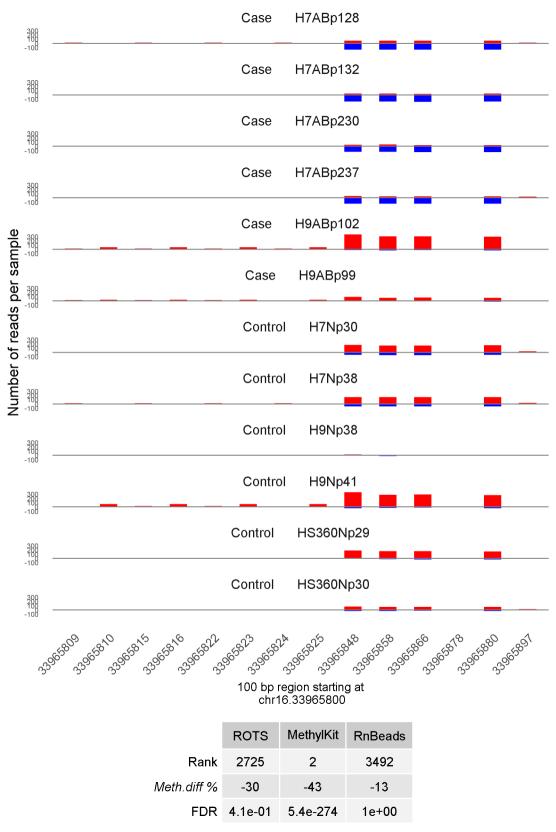
100 bp region starting at chr18.69467000

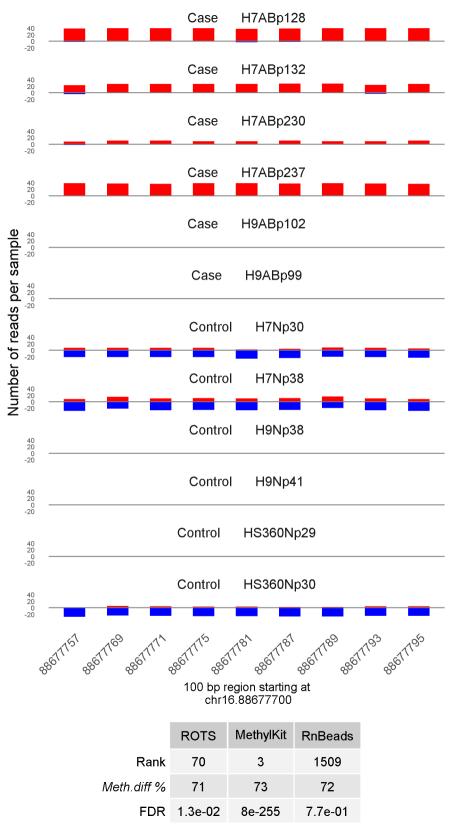
	ROTS	MethylKit	RnBeads
Rank	1029	2258	98
Meth.diff %	43	45	43
FDR	2e-01	2.6e-07	1.9e-02

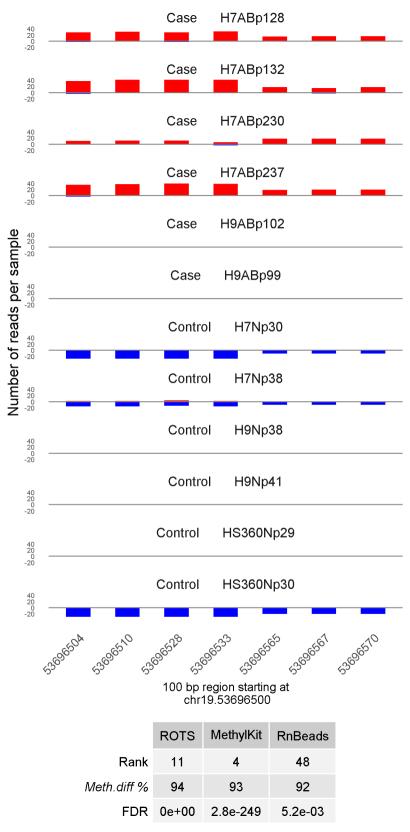


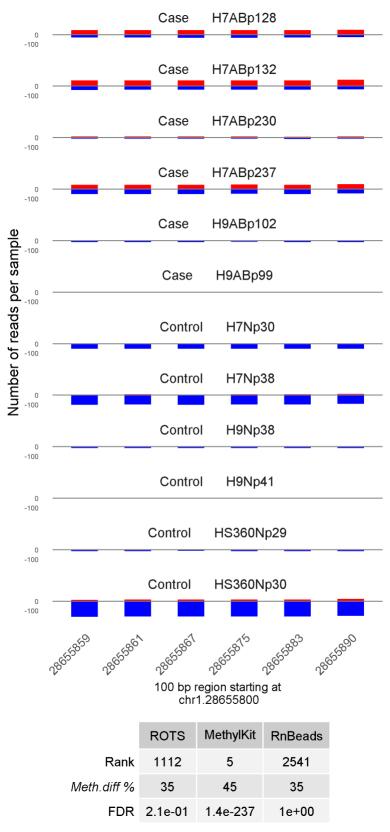


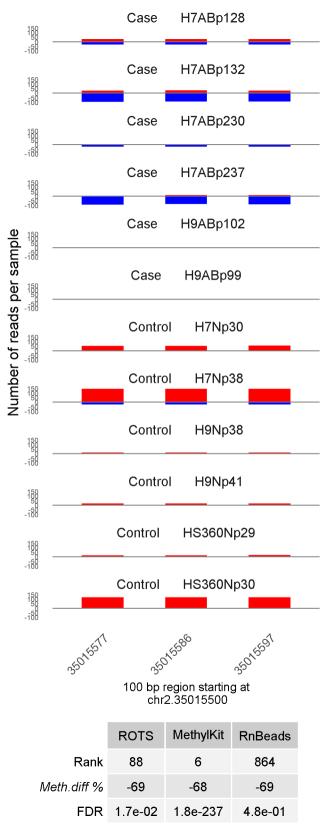


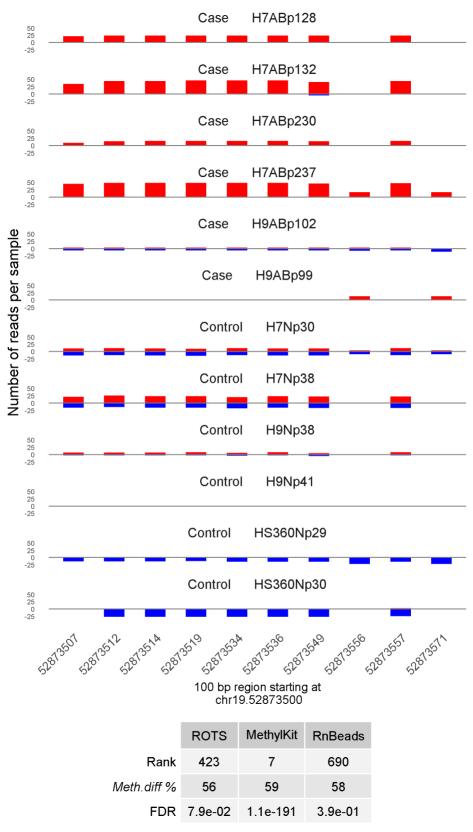


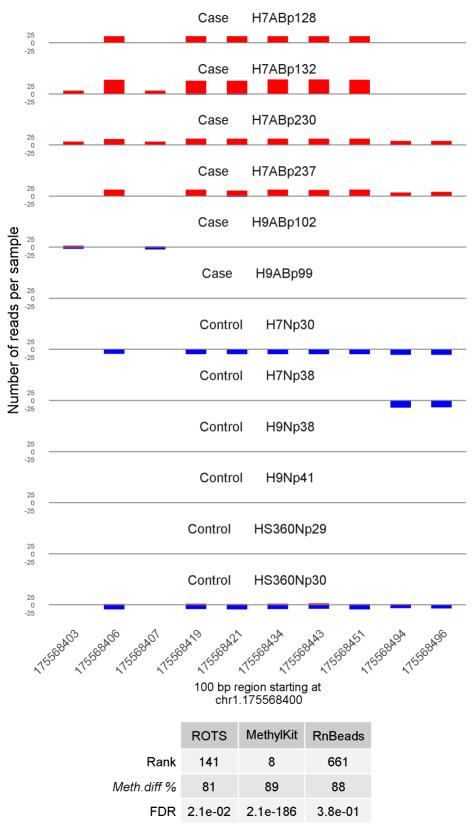


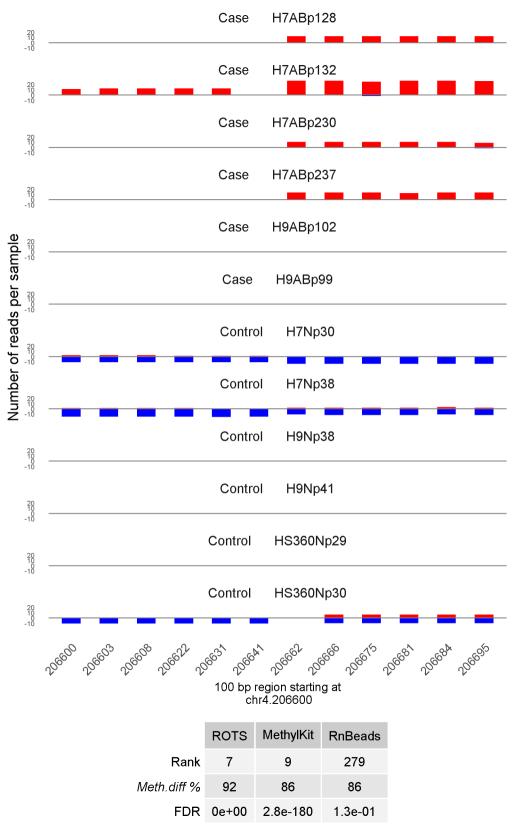


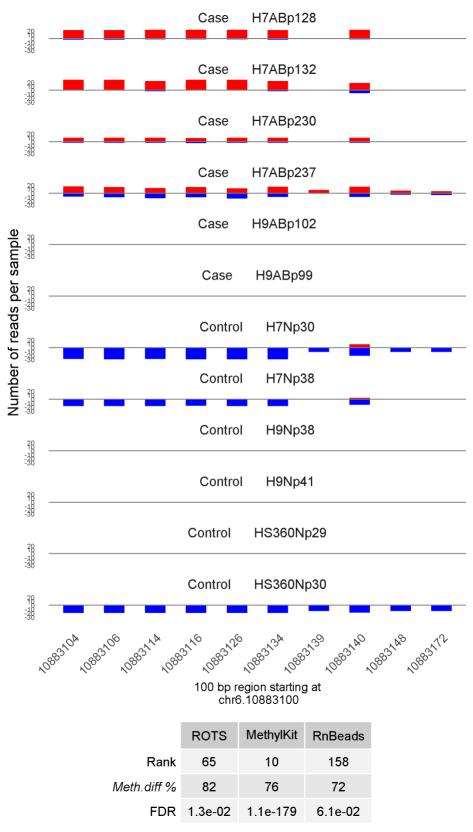


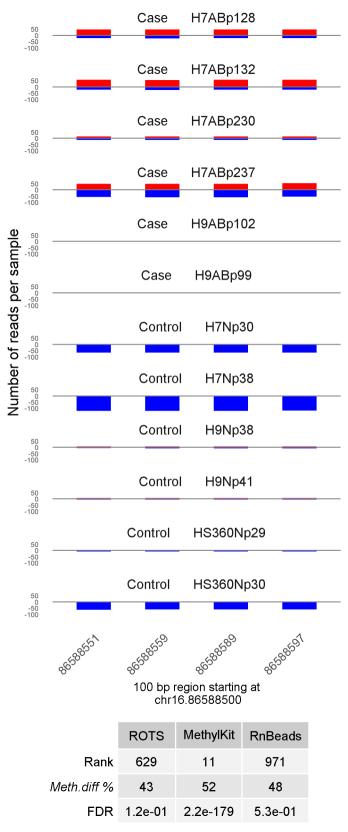


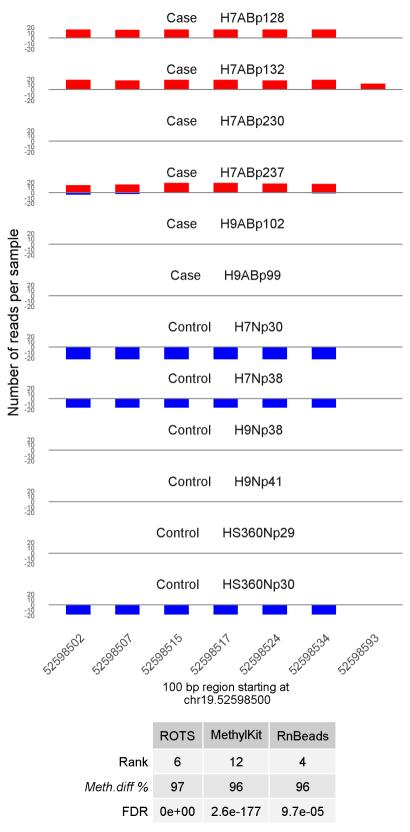


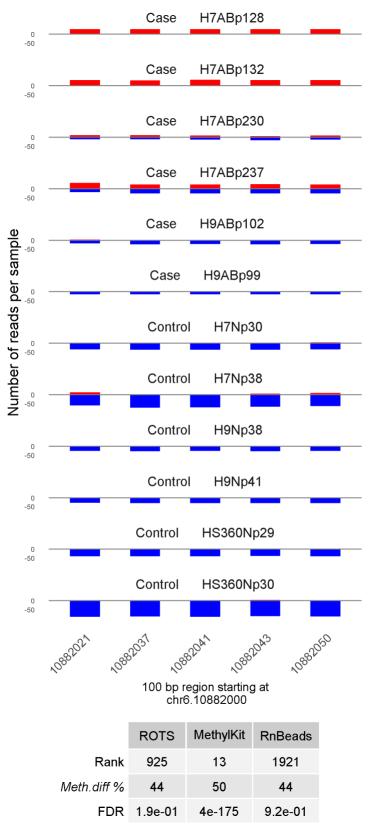


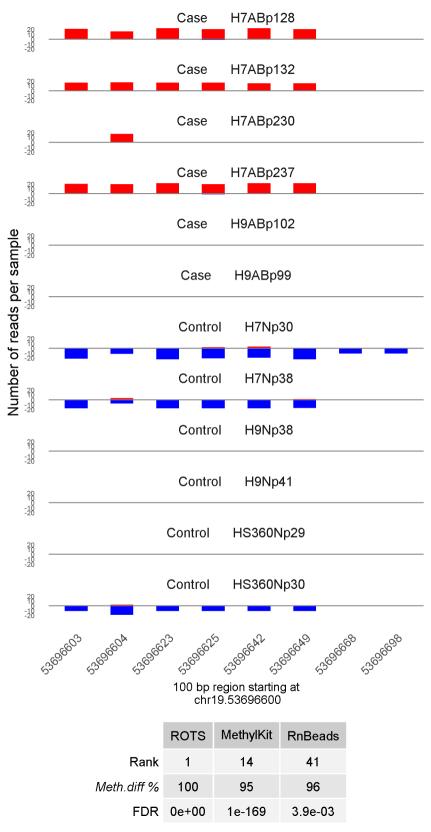


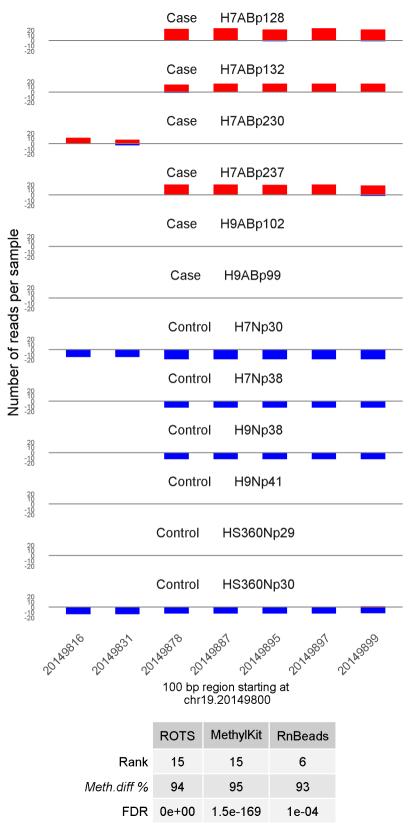


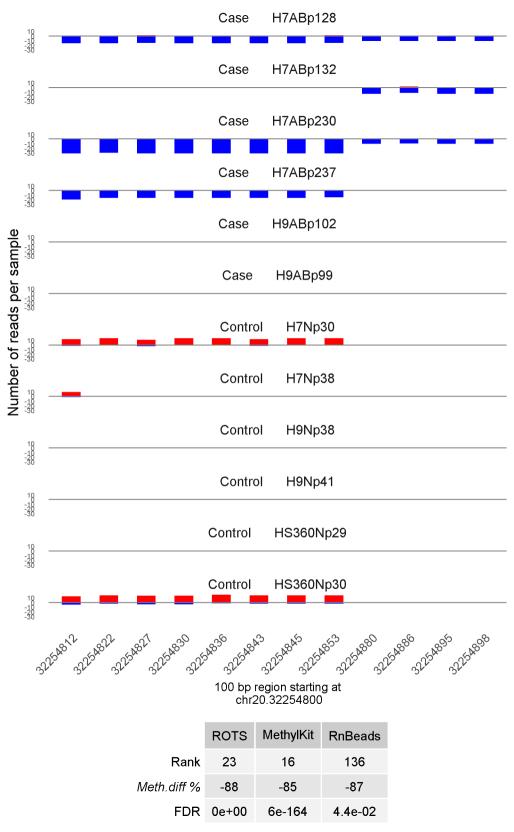


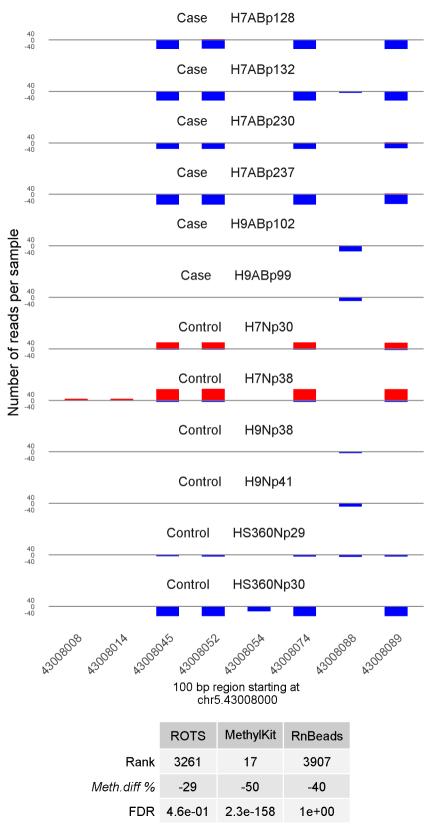


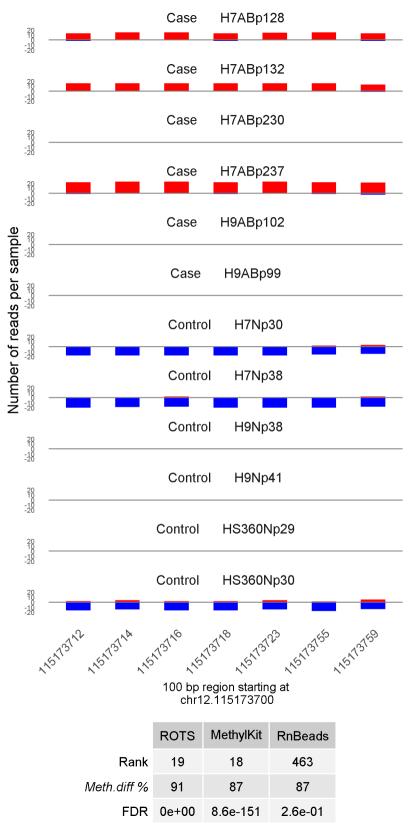


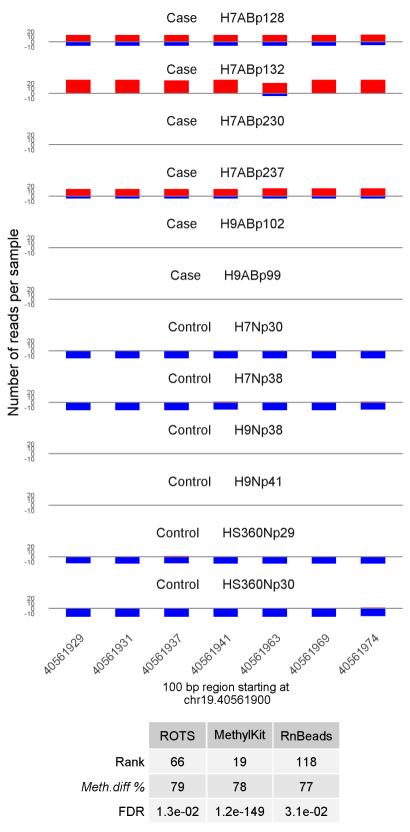


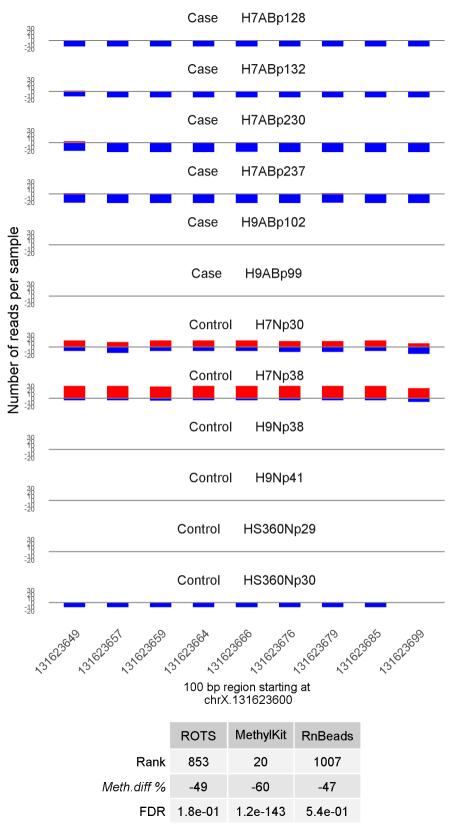


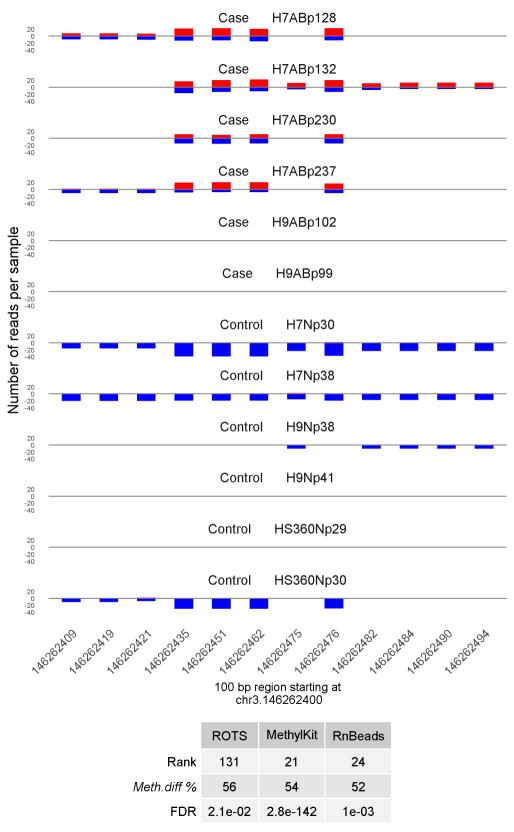


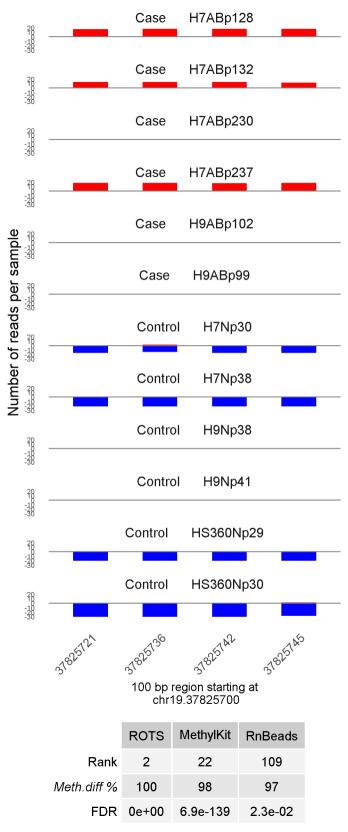


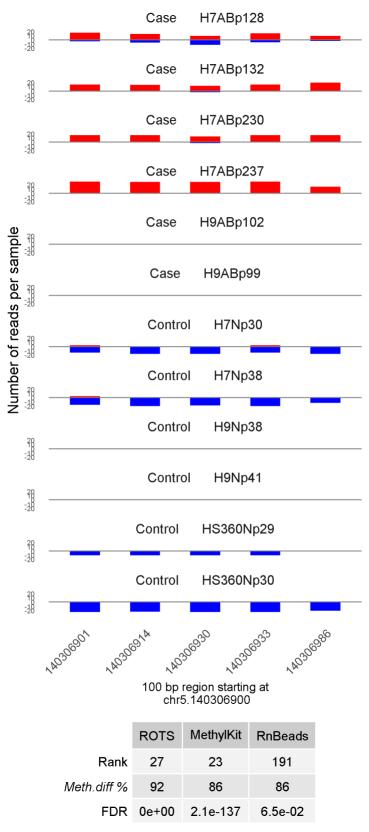


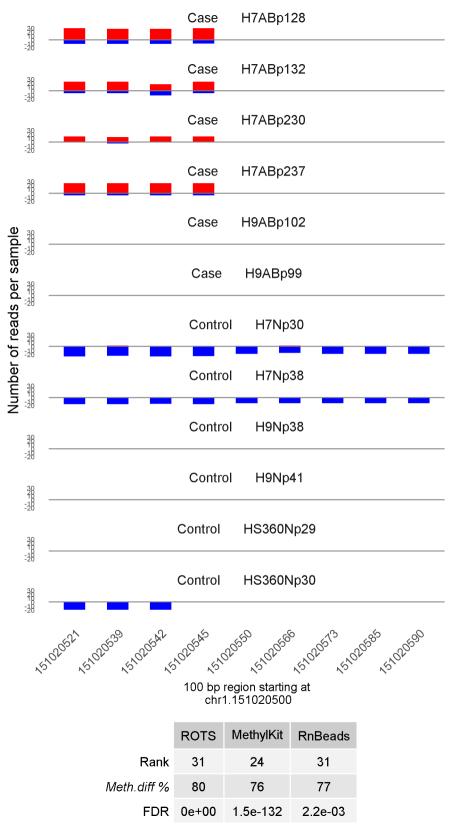


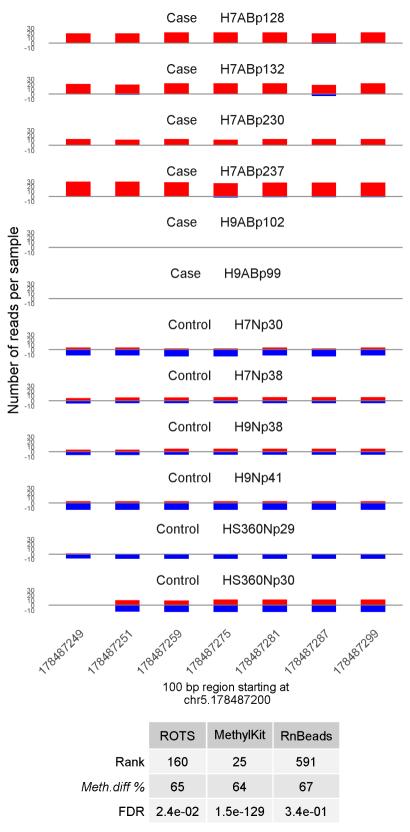


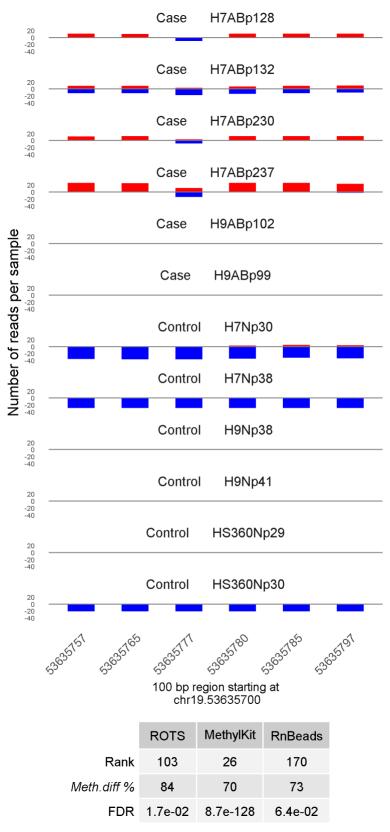


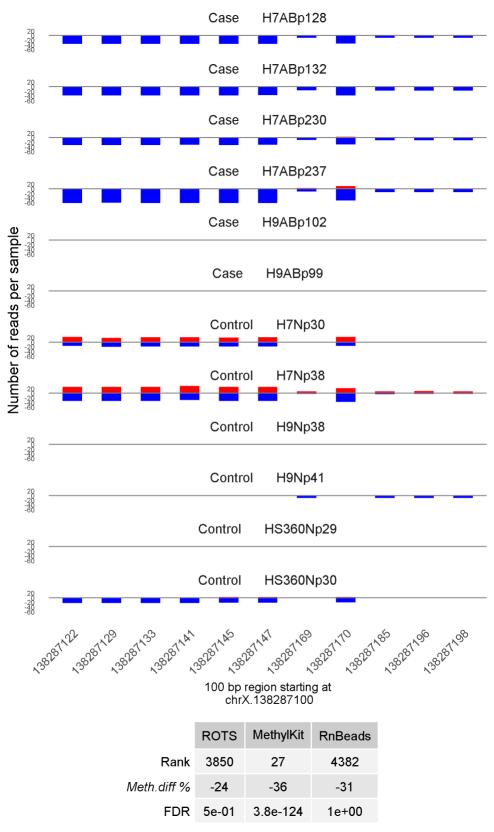


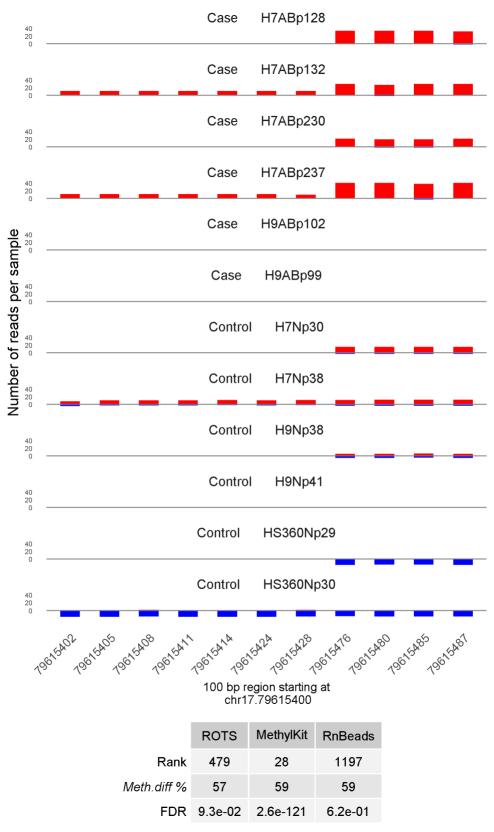


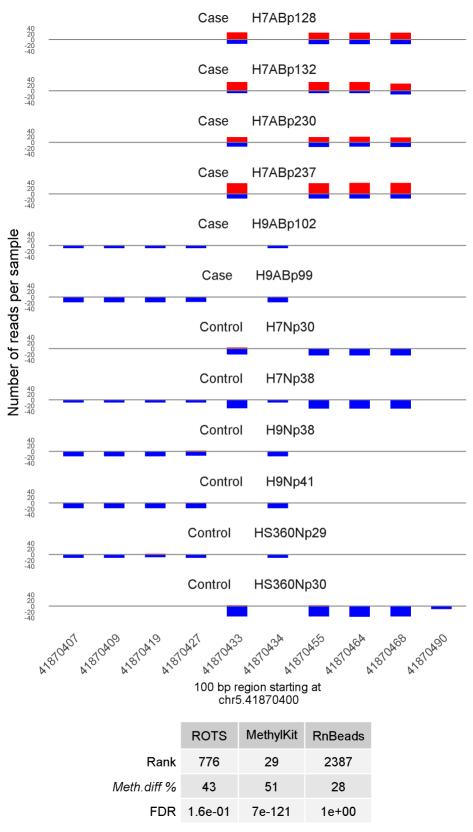


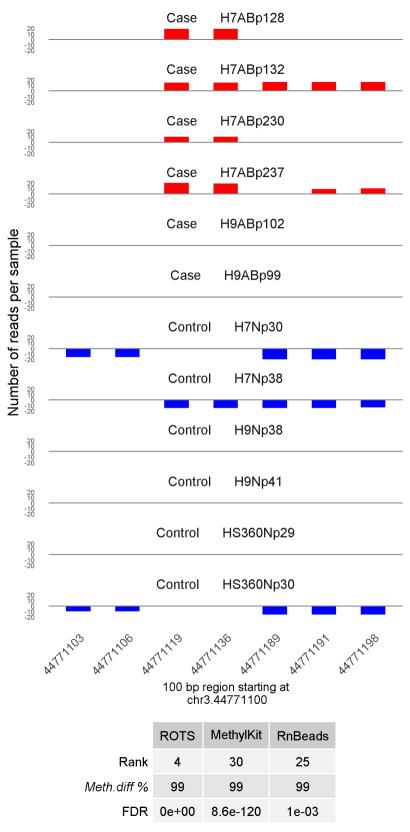


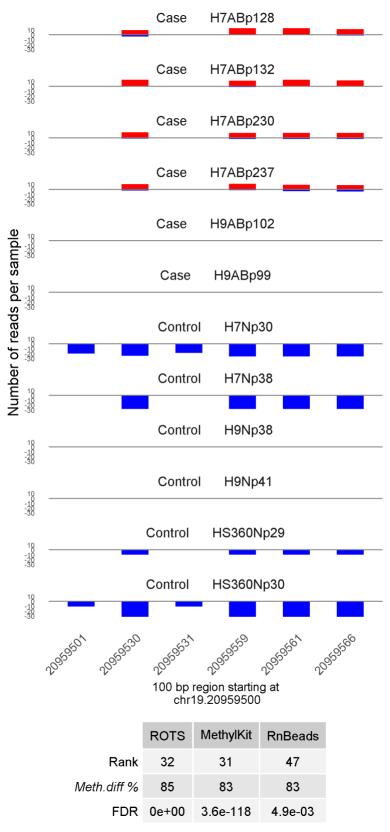


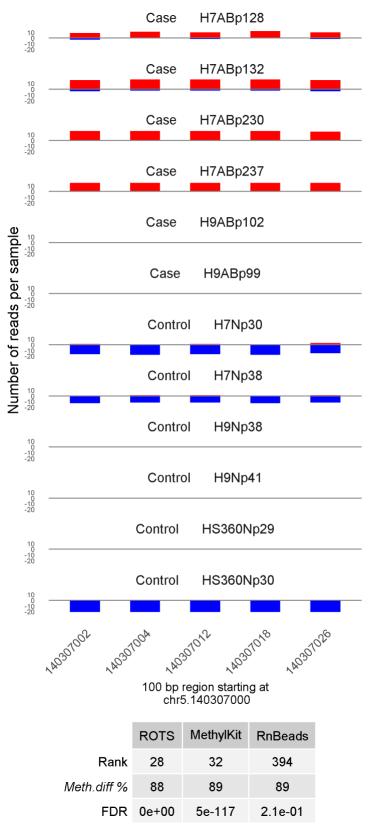


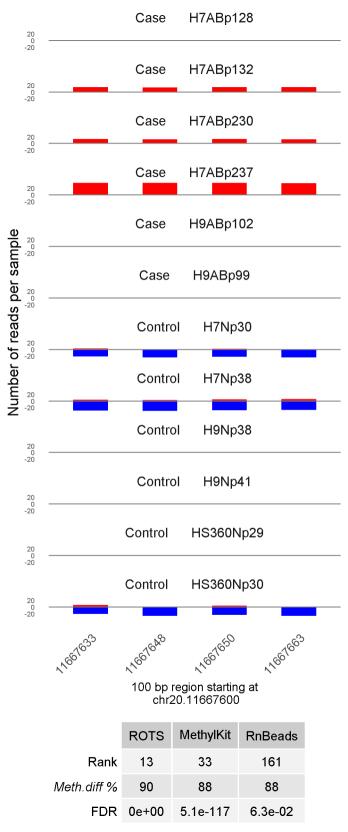


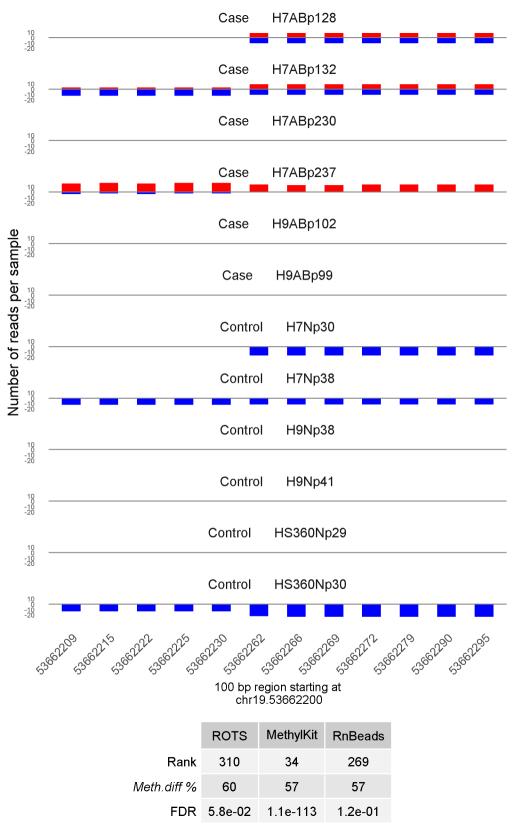


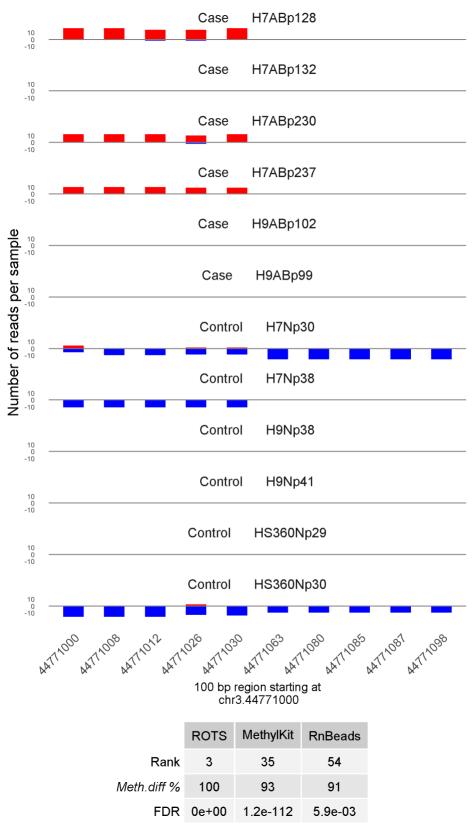


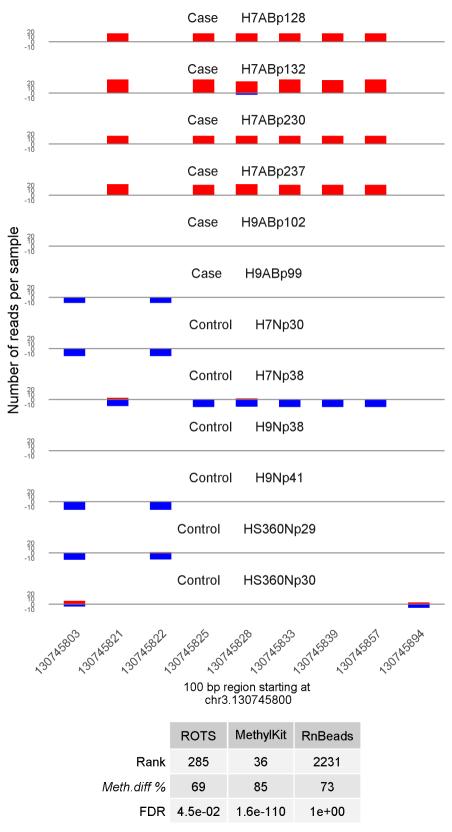


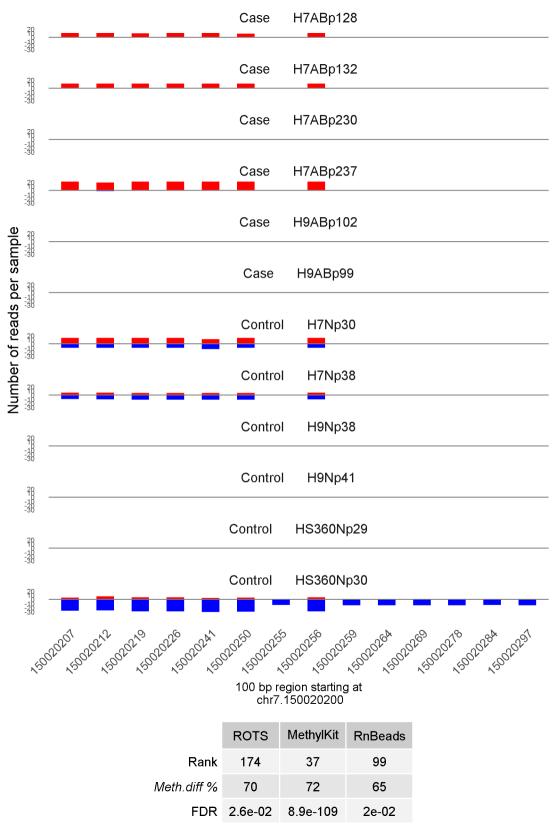


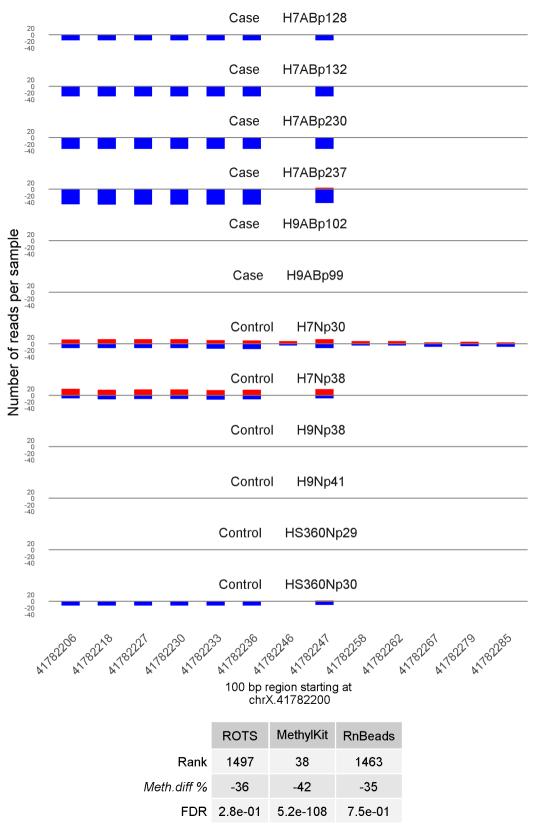


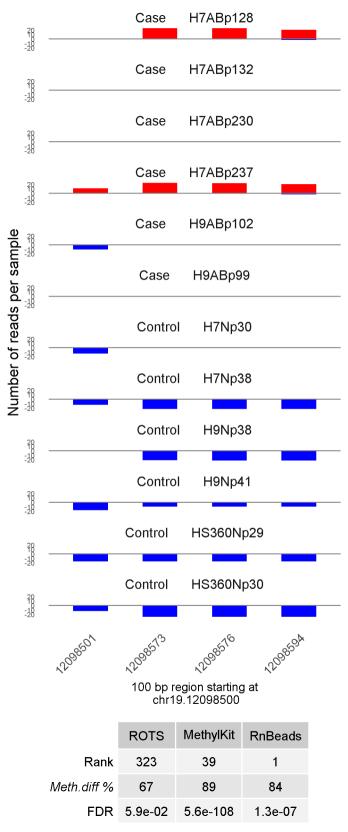


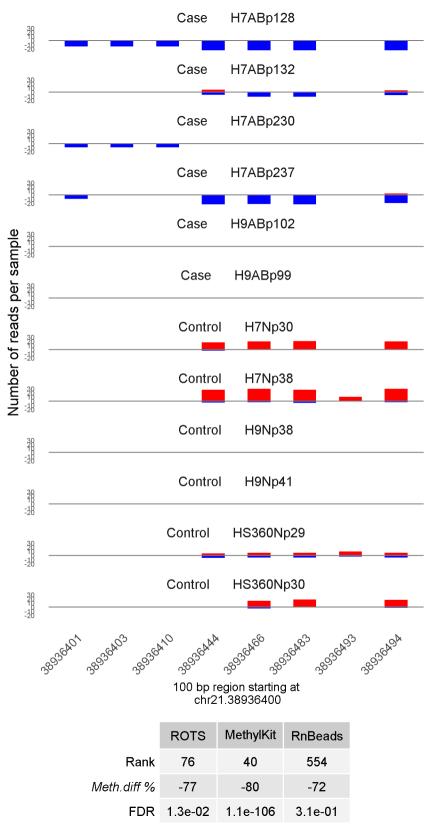


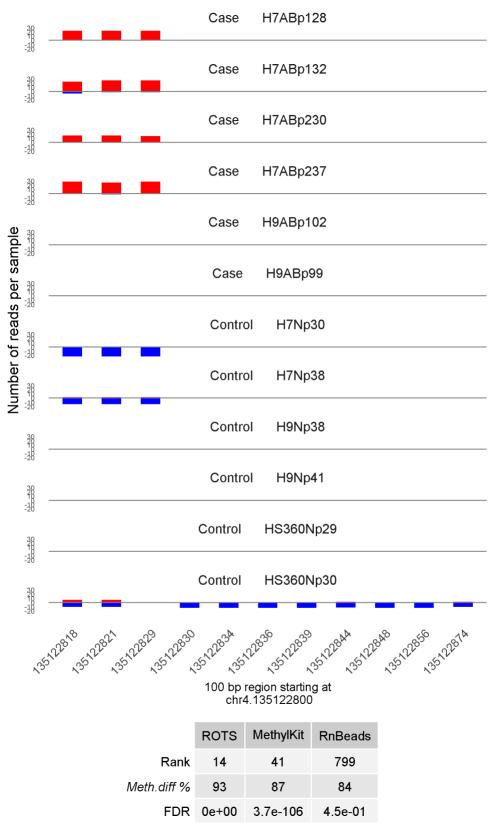


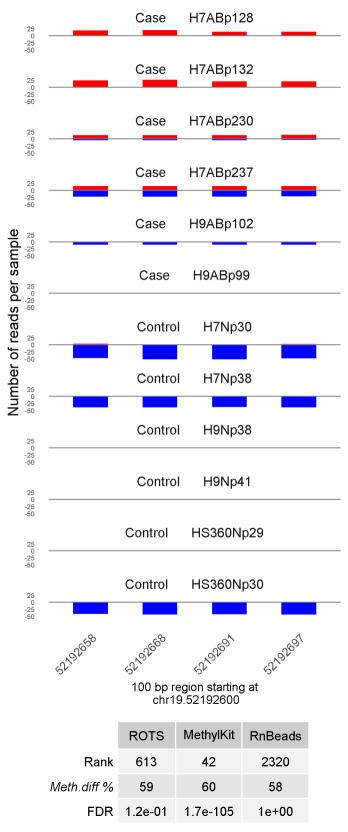


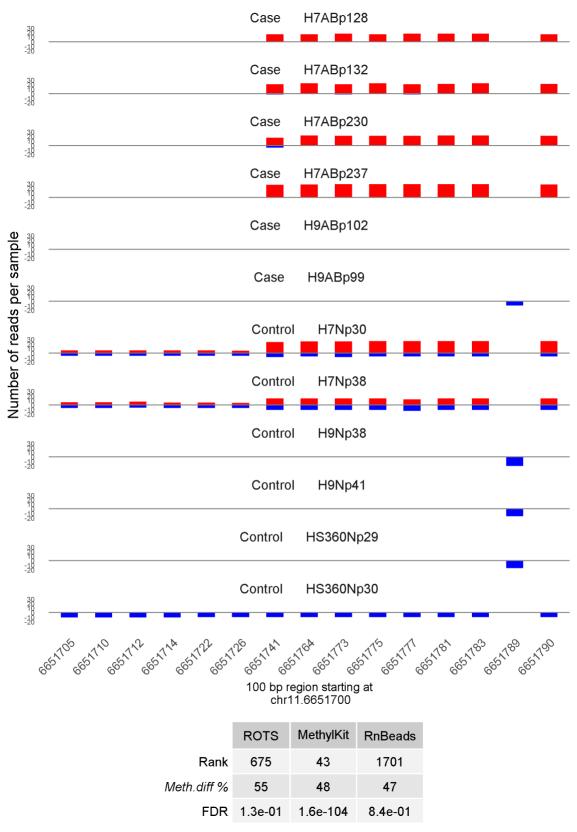


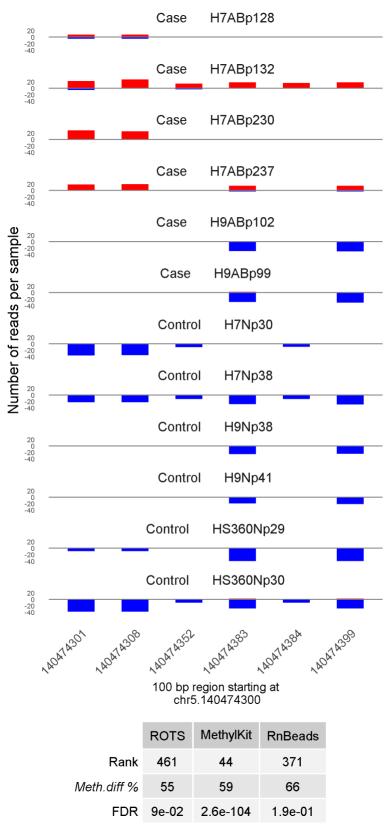


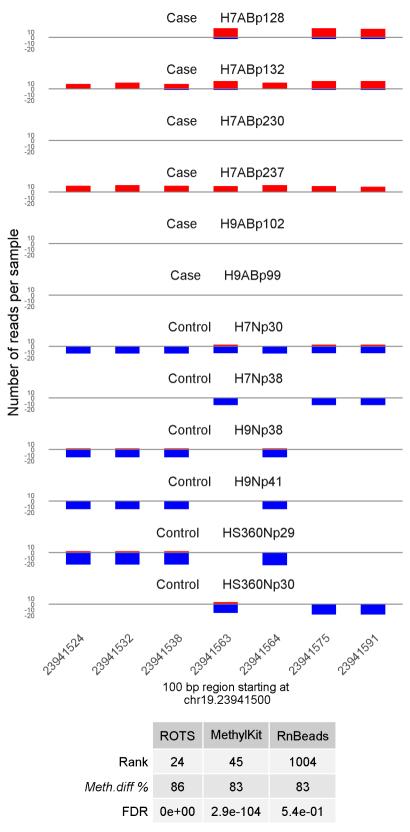


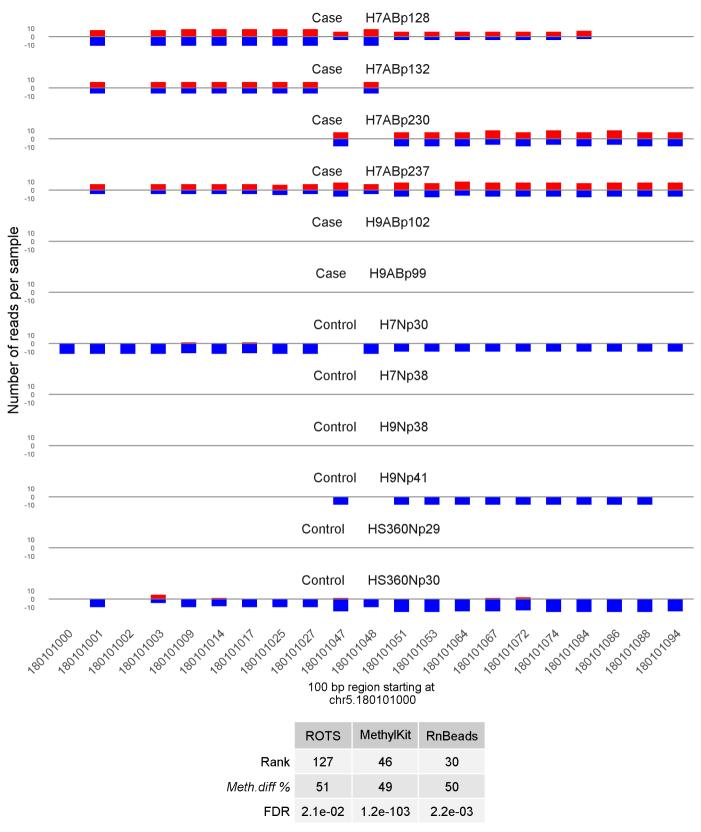


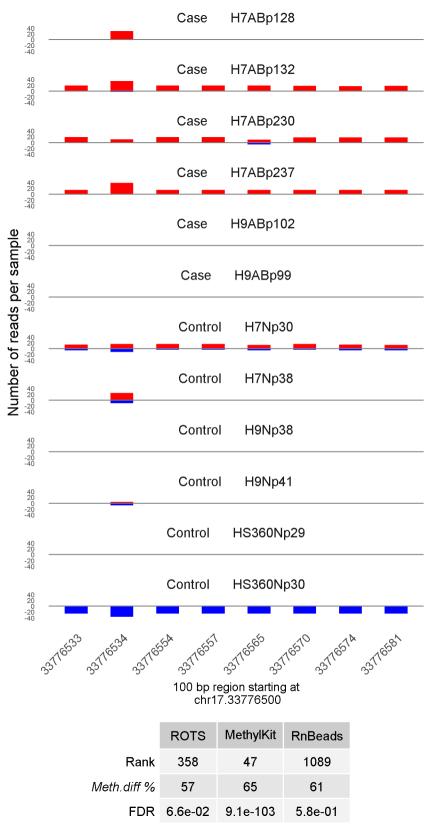


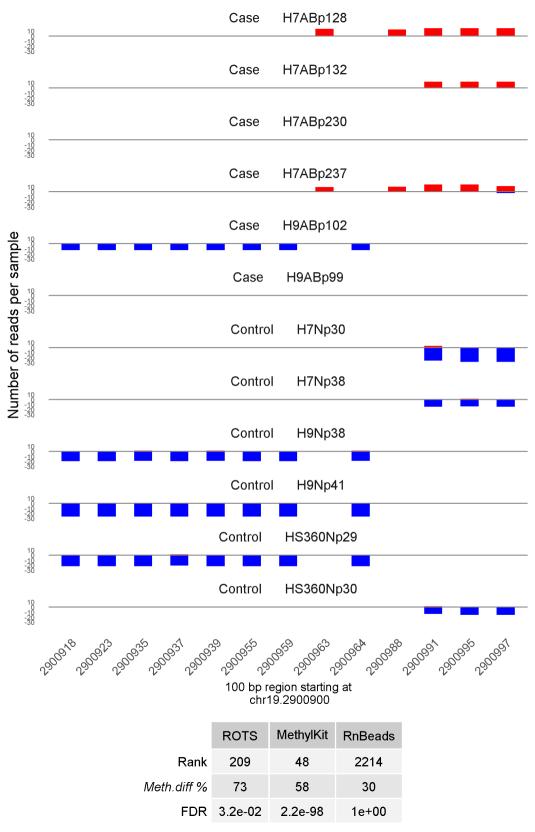


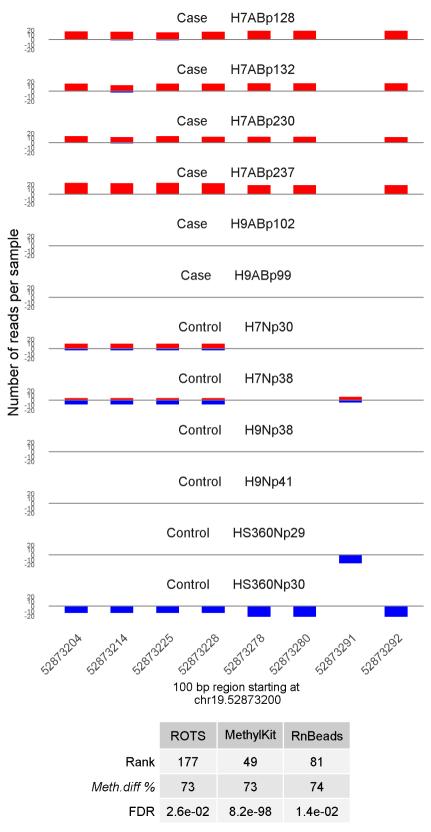


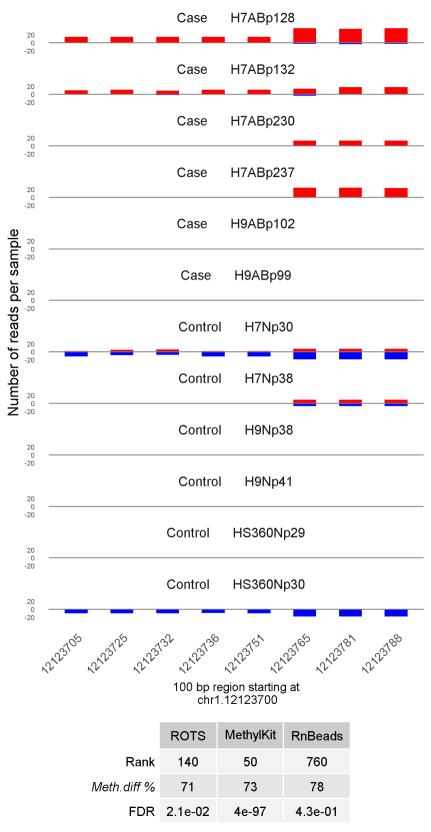


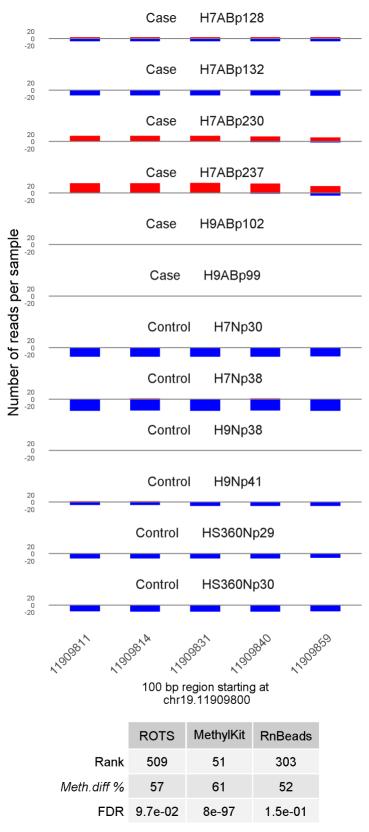


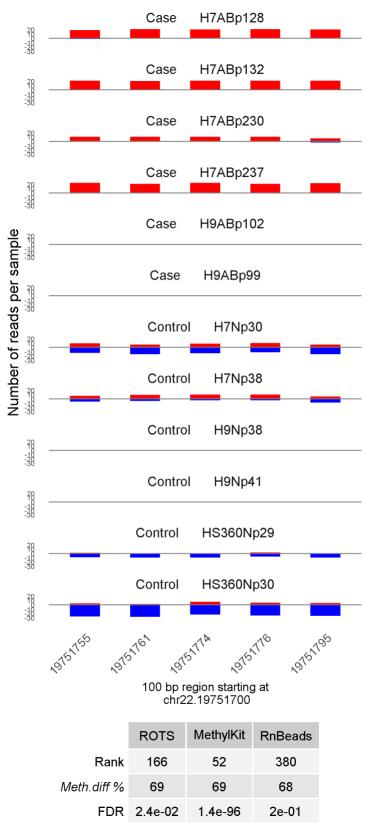


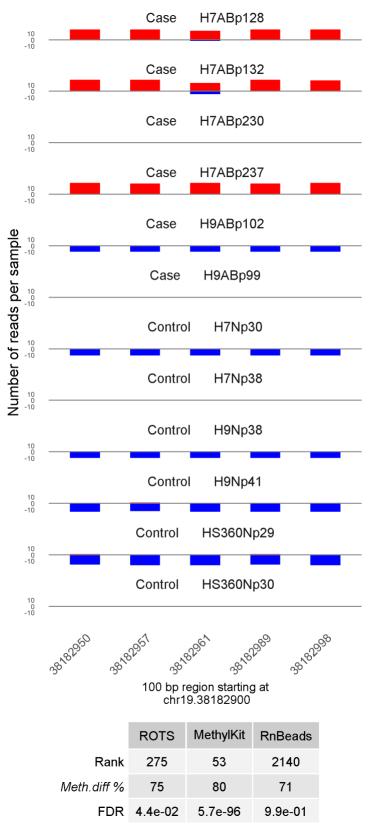


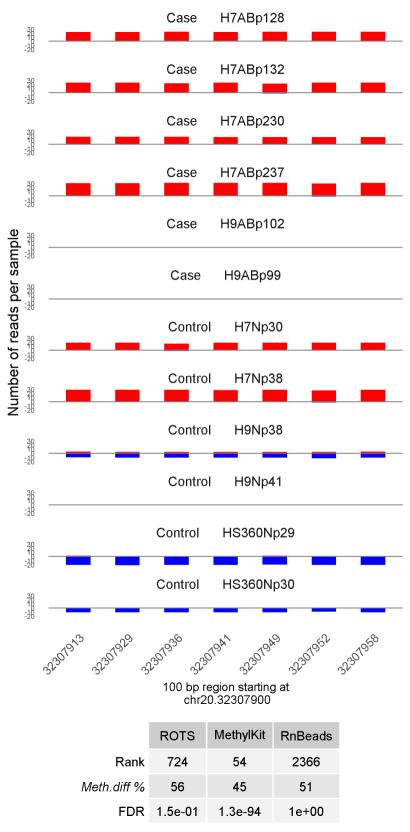


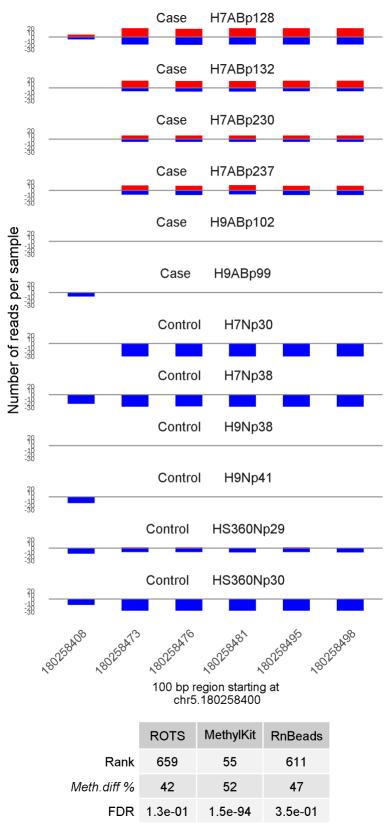


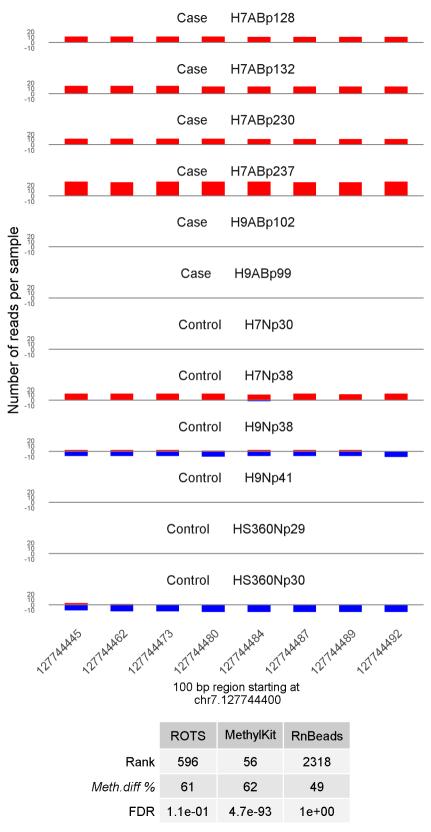


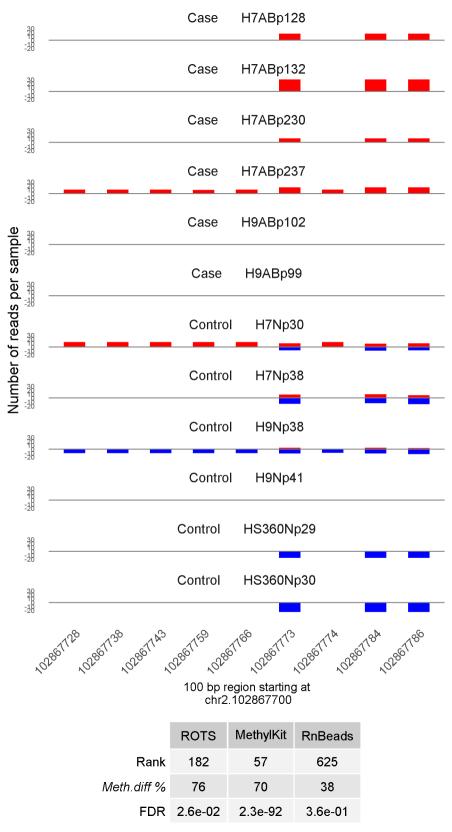


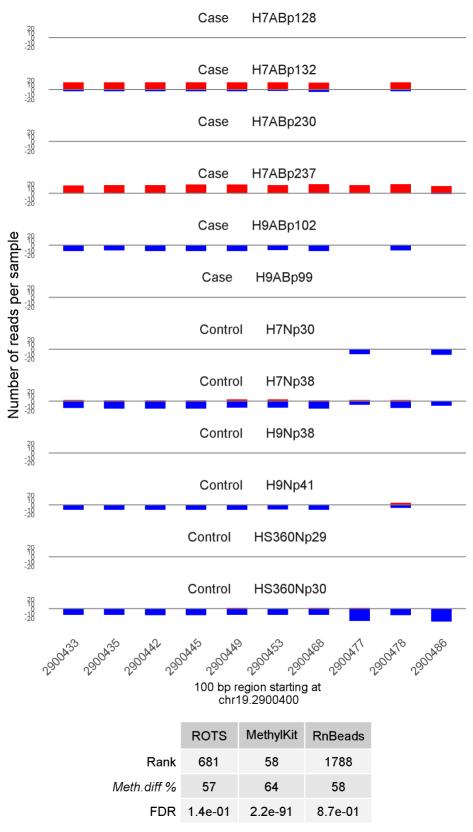


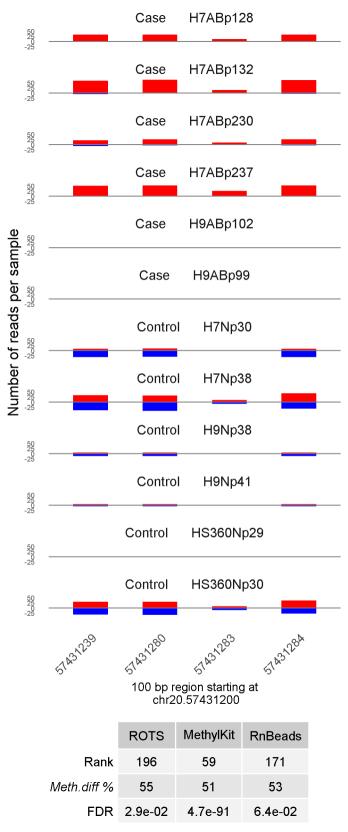


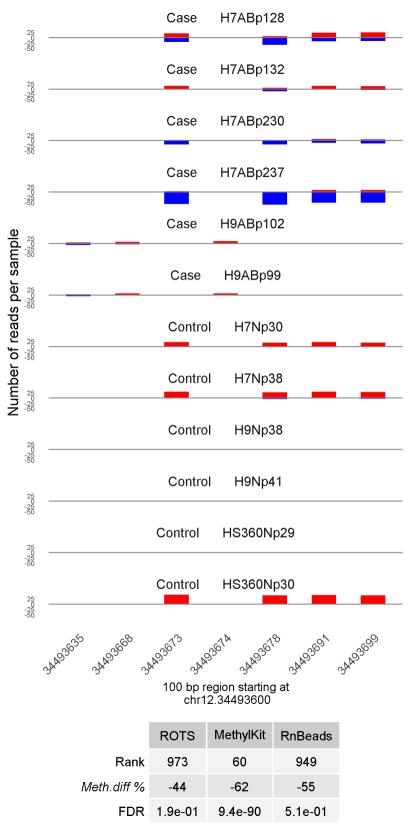


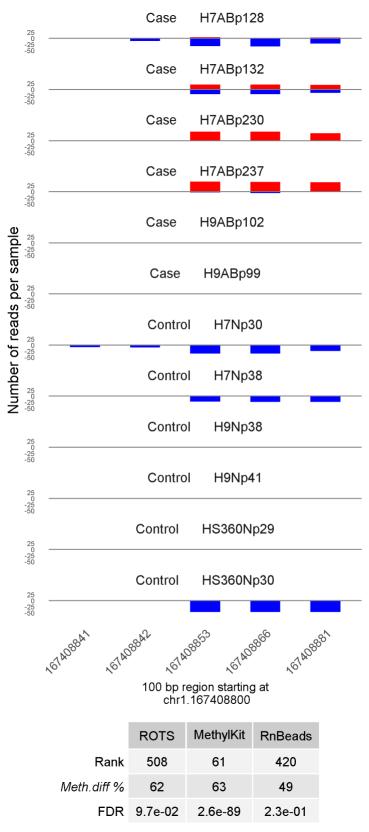


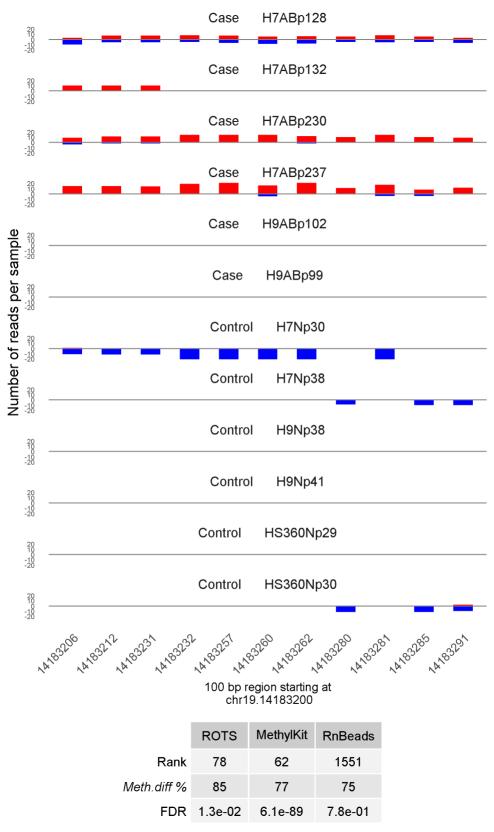


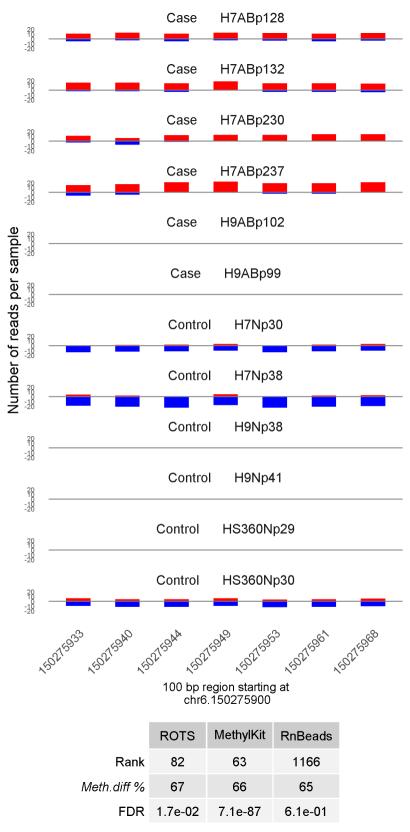


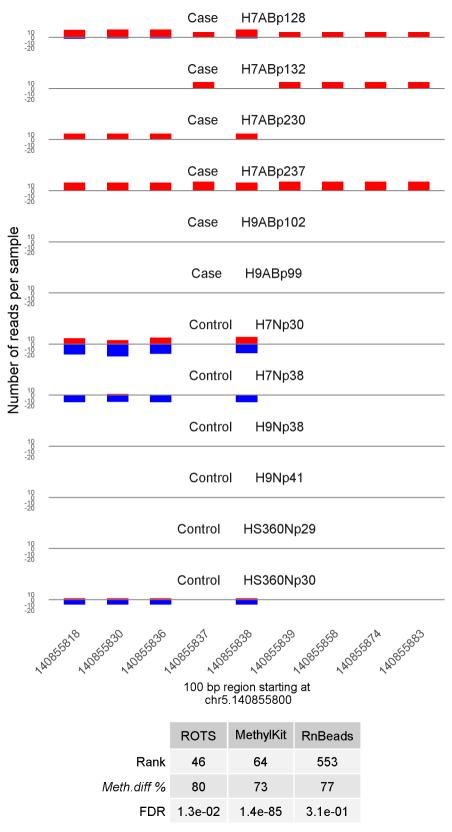


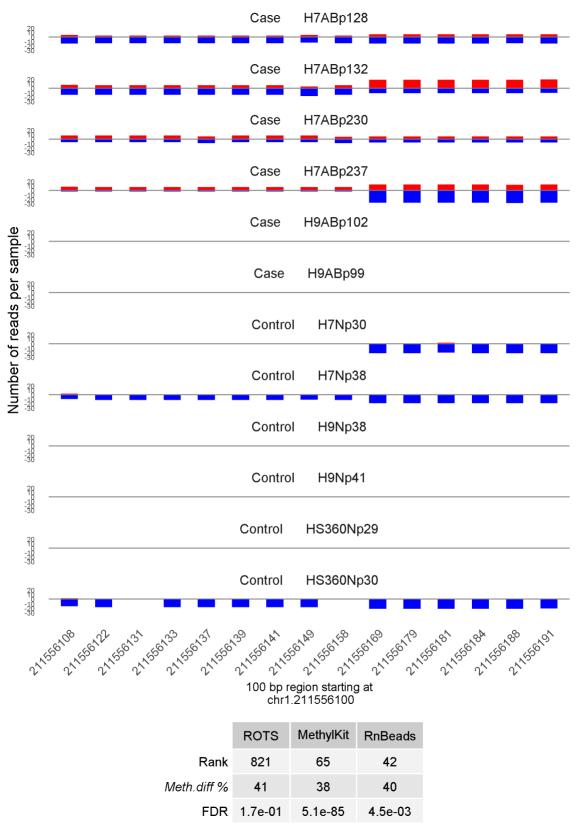


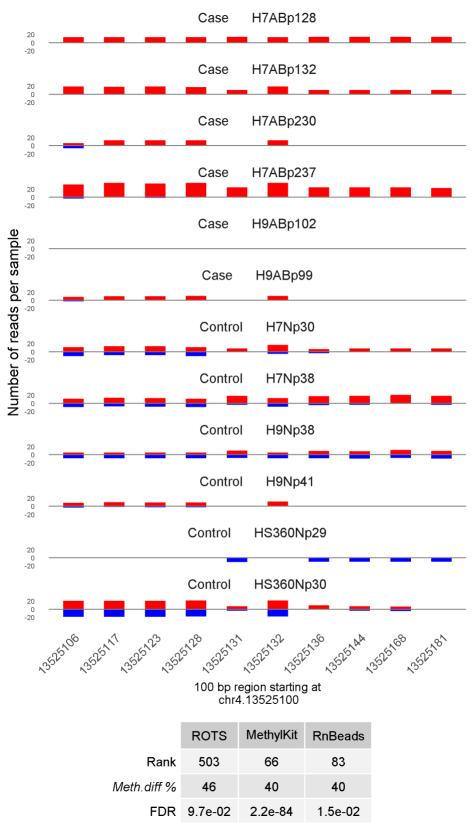


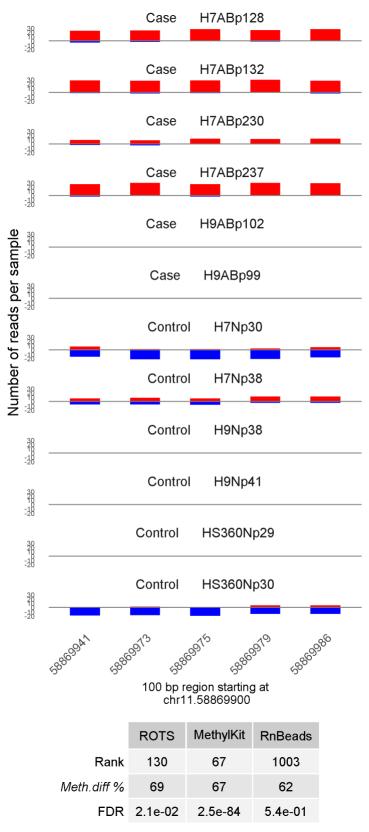


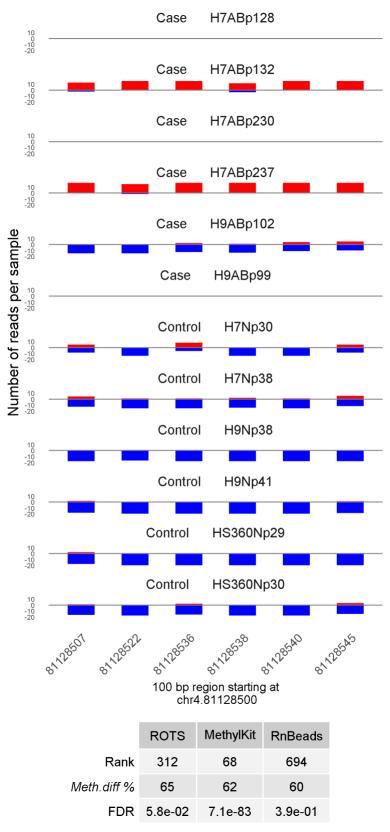


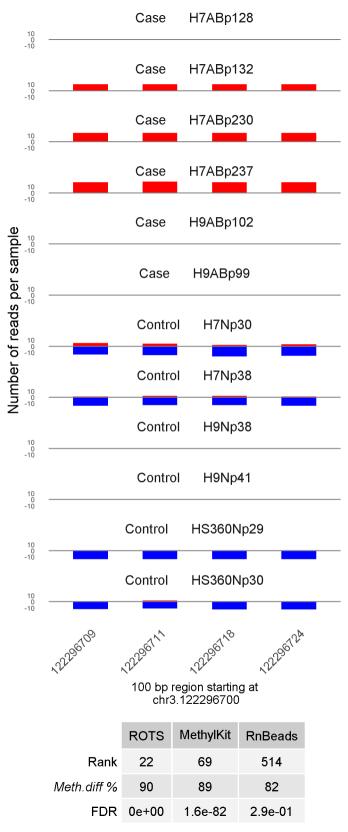


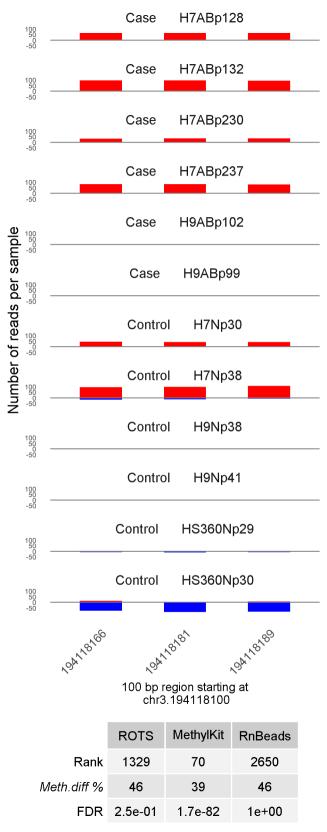


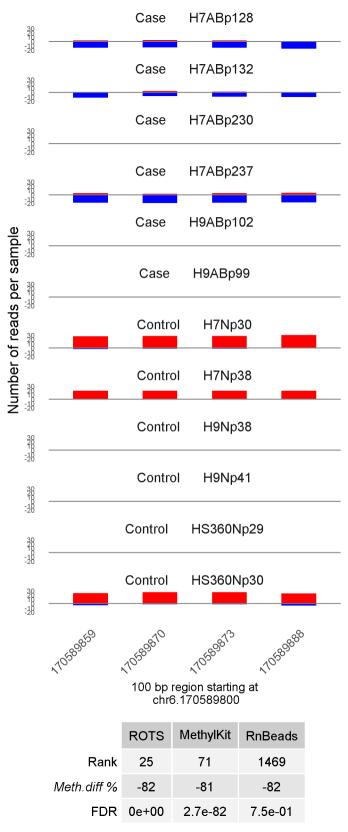


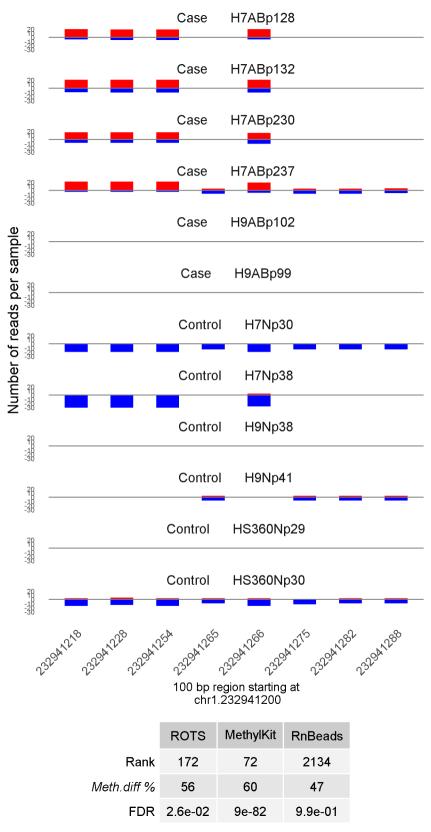


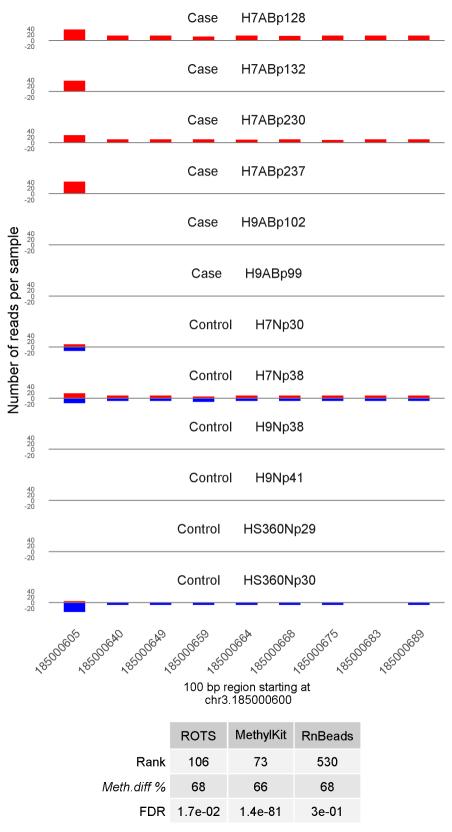


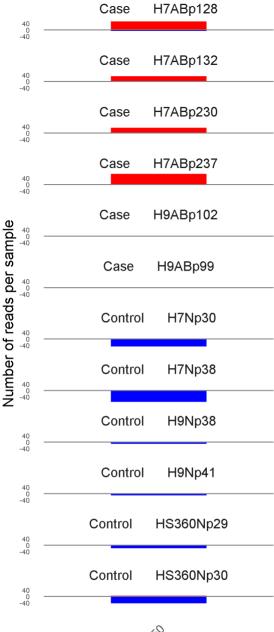














100 bp region starting at chr19.52598300

	ROTS	MethylKit	RnBeads
Rank	18	74	250
Meth.diff %	87	87	87
FDR	0e+00	1.8e-81	1.1e-01

