Τα γραφήματα δημιουργήθηκαν με τη συνάρτηση motifVisual, η οποία δέχεται ως εισόδους τον φάκελο στον οποίο θα αποθηκευτούν τα γραφήματα, τον πίνακα [data.plot3.4](#data_plot_34) και τη λίστα [report.list](#report_list), η οποία είναι το αποτέλεσμα του motif enrichment και έχει παραχθεί με τη συνάρτηση PWMEnrich.

1) Density plot-P values of Motifs

* διάγραμμα κατανομής των τιμών της p.value

2)Plots per TADs (φάκελος που για κάθε TAD περιέχει ένα φάκελο με διαγράμματα για τα TFs που βρέθηκαν στο συγκεκριμένο TAD) – δεδομένα εισόδου: report.list

* PWMEnrich\_Image
  + εικόνα με τα top 10 μοτίβα που βρέθηκαν στο TAD και πληροφορίες σχετικά με το TF, στο οποίο αντιστοιχούν
  + δημιουργία με τη συνάρτηση plot(), του πακέτου PWMEnrich.
* Histogram Image
  + επιλογή των γραμμών με μοτίβα που έχουν αντιστοιχηθεί σε TF (target)
  + επιλογή των στηλών: target, tad, p.value
  + ομαδοποίηση ανά target, tad και summarise, ώστε p.value = mean(p.value) σε περίπτωση που περισσότερα του ενός μοτίβα αντιστοιχούν στον ίδιο TF
  + αν οι γραμμές του πίνακα ξεπερνούν τις 30, τότε επιλογή των 30 πρώτων γραμμών και αναφορά αυτής της επιλογής ως υπότιτλο στο γράφημα
  + δημιουργία ιστογράμματος, στο οποίο ο άξονας x αντιπροσωπεύει τη μέση τιμή p.value, ο άξονας y τους TFs.

3)Transcription Factors (φάκελος που για κάθε TF περιέχει ένα φάκελο με διαγράμματα για τα TADs, στα βρέθηκε και τα μοτίβα πρόσδεσης του πάνω στα οποία αντιστοιχήθηκε από τις ακολουθίες εισόδου) - δεδομένα εισόδου: report.list, data.plot3.4

* Barplot
  + διάσπαση της στήλης P.value του πίνακα data,plot3.4
  + μετατροπή της στήλης P.value σε numeric type
  + ομαδοποίηση του πίνακα ανά TFs, TAD και summarise, ώστε P.value = mean(P.value)
  + ομαδοποίηση ανά TFs
  + για κάθε ομάδα TFs ταξινόμηση κατά φθίνουσα τιμή P.value
  + αν οι γραμμές του πίνακα ξεπερνούν τις 30, τότε επιλογή των 30 πρώτων γραμμών και αναφορά αυτής της επιλογής ως υπότιτλο στο γράφημα
  + δημιουργία ιστογράμματος, στο οποίο ο άξονας x αντιπροσωπεύει τη μέση τιμή P.value, ο άξονας y τα TADs.
* Motifs
  + με αρχή τη λίστα report.list, δημιουργία της λίστα PFMs, κάθε αντικείμενο της οποίας αντιστοιχεί στον Position Frequency Matrix ενός μοτίβου πρόσδεσης που βρέθηκε στις ακολουθίες και έχει όνομα το id του μοτίβου
  + για κάθε ομάδα TFs (από την προηγούμενη ομαδοποίηση) διάσπαση της στήλης Motifs
  + αντιστοίχιση κάθε μοτίβου με τον Position Frequency Matrix του από τη λίστα PFMs
  + δημιουργία εικόνας με τα sequence logos όλων των μοτίβων

4)TFs in different TADs

* ομαδοποίηση του πίνακα data.plot3.4 ανά TFs, TAD,Motifs και summarise ώστε P.value = mean(P.value)
* μέτρηση το πόσες φορές (n) εμφανίζεται κάθε TF στον πίνακα data.plot3.4
* συνένωση του αποτελέσματος με τον πίνακα data.plot3.4
* ταξινόμηση του πίνακα κατά φθίνουσα τιμή της στήλης n
* αν οι γραμμές του πίνακα ξεπερνούν τις 30, τότε επιλογή των 30 πρώτων γραμμών
* δημιουργία ιστογράμματος, στο οποίο ο άξονας x αντιπροσωπεύει τον αριθμό των TADs, στα οποία βρέθηκε κάθε TF και ο άξονας y τους TFs.

5) Top Transcription Factors

* χρήση του ίδιου πίνακα από το προηγούμενο διάγραμμα
* δημιουργία ιστογράμματος, στο οποίο ο άξονας x αντιπροσωπεύει τη μέση τιμή P.value, ο άξονας y τους TFs και το χρώμα των ράβδων την τιμή της στήλης n
* το ιστόγραμμα είναι ταξινομημένο κατά φθίνουσα τιμή της στήλης n

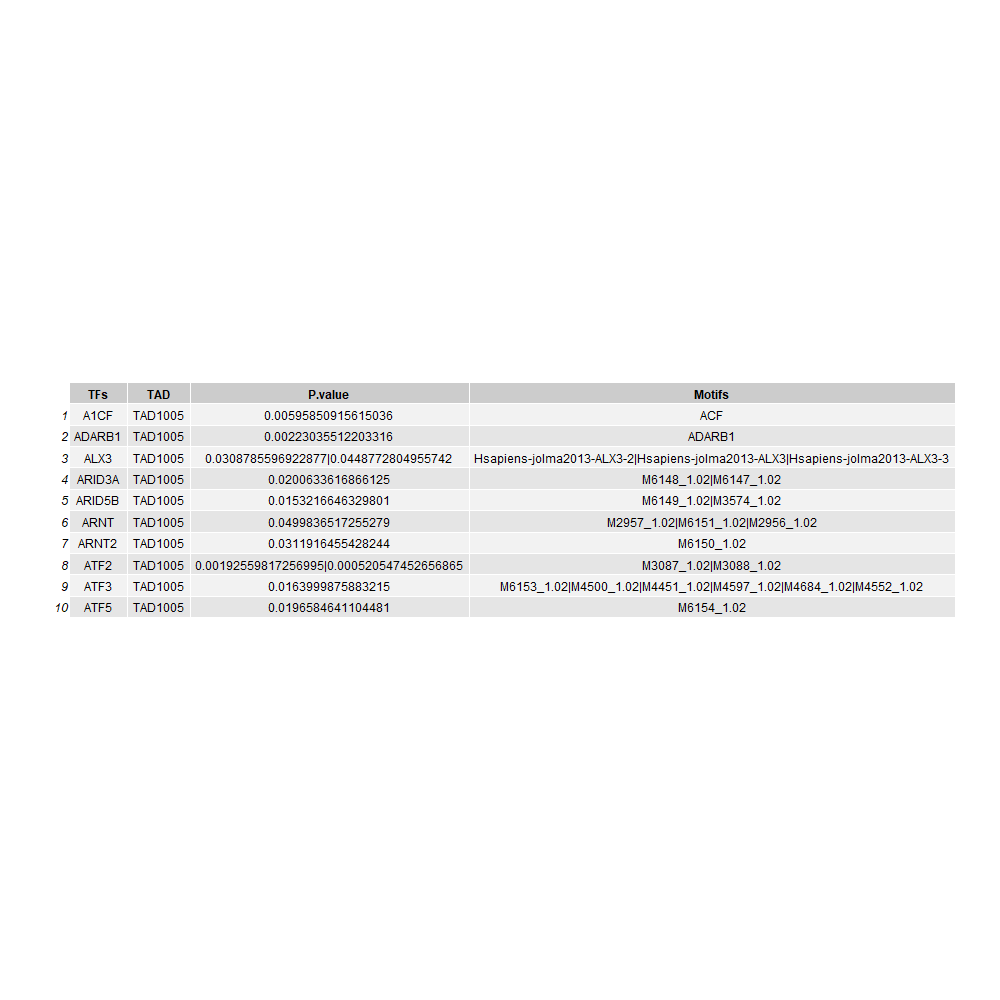
6) Top Transcription Factors-2

* χρήση του ίδιου πίνακα από το προηγούμενο διάγραμμα
* δημιουργία ιστογράμματος, στο οποίο ο άξονας x αντιπροσωπεύει τη μέση τιμή P.value, ο άξονας y τους TFs και το χρώμα των ράβδων την τιμή της στήλης n
* το ιστόγραμμα είναι ταξινομημένο κατά φθίνουσα τιμή της μέσης P.value

**Βήματα υπολογισμού της λίστας report.list**

* χρήση του πίνακα εισόδου (integrated\_table\_with\_sign\_tads.csv / έξοδος της TADiff.R)
* επιλογή των events, τα οποία στη στήλη Gene\_locus βρίσκεται κάποια από τις τιμές promoter|intergenic|fiveUTR
* επιμήκυνση των ακολουθιών με μήκος <30 κατά ±25
* συνένωση των ακολουθιών ανά TAD
* για κάθε ακολουθία :
  + get request Ensembl Rest Api με τις χρωμοσωμικές συντεταγμένες του για απόκτηση της ακολουθίας βάσεων
  + αποθήκευση της ακολουθίας σε αρχείο fasta (ένα ενιαίο αρχείο με όλες τις ακολουθίες)
* για κάθε TAD αποθήκευση σε πίνακα την πρώτη και την τελευταία σειρά του αρχείου fasta που αντιστοιχεί στις ακολουθίες του
* εκτέλεση motif enrichment ανά TAD και αποθήκευση των αποτελεσμάτων στη λίστα report.list

**Βήματα υπολογισμού του πίνακα data.plot3.4**

****

Στιγμιότυπο του πίνακα data.plot3.4

* για κάθε αντικείμενο της λίστας report.list:
  + αφαίρεση των μοτίβων “UW.Motif.####”
  + επιλογή των στηλών target,tad,p.value,id
  + ομαδοποίηση κατά target, tad και summarize P.value , Motifs με collapse = "|"
  + ένωση σε ένα πίνακα (rbind) με τα αποτελέσματα από τα υπόλοιπα αντικείμενα της λίστας.
* μετονομασία των στηλών σε "TFs", "TAD", "P.value", "Motifs"
* επιλογή των στηλών TFs, TAD, P.value και αποθήκευση στον πίνακα data.plot3.4
* στον προηγούμενο πίνακα ομαδοποίηση ανά TFs και summarise τις υπόλοιπες στήλες με collapse = "|"
* έπειτα για κάθε γραμμή/TFs
  + επιλογή της τιμής της στήλης Motifs
  + διάσπαση ως προς “|”
  + επιλογή μοναδικών στοιχείων/μοτίβων
  + ένωση πάλι με collapse = "|"
  + επανατοποθέτηση στη στήλη Motifs
* επιλογή των στηλών TFs, Motifs του πίνακα
* συνένωση του τελευταίου πίνακα με τον πίνακα data.plot.3.4, ως προς την κοινή στήλη TFs