Spalten mir181 BS tabelle

1-5: "seqnames","start" ,"end","width","strand"

🡪 Position der Bindestelle

6-8: "scoreSum" "scoreMean" "scoreMax"

🡪 score der Bindestelle, berechnet aus den purclips scores der Bindestelle, entweder die Summe der Mean oder das Maximum aller pureclipscores innerhalb der Bindestelle

9-12: "geneType" "geneName" "geneID" "region"

🡪 Infos zu dem zugeordneten Gen

13: mir\_IP

🡪 kommt aus früherem script unwichtig, kann gelöscht werden

14-18: "n\_mir181" "n\_mir181a" "n\_mir181b" "n\_mir181c" "n\_mir181d"

🡪 Anzahl der chimeric reads in der bindestelle oder in einem umkreis von 10nt vor oder hinter der bindestelle für die jeweilige miR (mir181 = alle, mir181a = nur 181a etc)

19: set

🡪 aus welcher Bindestellen definition die bindestelle kommt. Entweder ago\_bs\_mir181\_chi (= Ago bindestellen mit mind. 2 mir181 chimeras im Umkreis von 10nt) oder mir181\_enriched (bindestellen aus IP of IP) oder beides

20: "mir181BS\_ID"

🡪 ID to recognise this binding site

21-60: "WT" "KO" "geneID.2" "geneName.1"

[25] "region.1" "counts.bs.1\_KO" "counts.bs.2\_KO" "counts.bs.3\_KO"

[29] "counts.bs.4\_WT" "counts.bs.5\_WT" "counts.bs.6\_WT" "geneID.1"

[33] "counts.bg.1\_KO" "counts.bg.2\_KO" "counts.bg.3\_KO" "counts.bg.4\_WT"

[37] "counts.bg.5\_WT" "counts.bg.6\_WT" "resBs.baseMean" "resBs.log2FoldChange"

[41] "resBs.lfcSE" "resBs.stat" "resBs.pvalue" "resBs.padj"

[45] "resBg.baseMean" "resBg.log2FoldChange" "resBg.lfcSE" "resBg.stat"

[49] "resBg.pvalue" "resBg.padj" "tpm.counts.bg.1\_KO" "tpm.counts.bg.2\_KO"

[53] "tpm.counts.bg.3\_KO" "tpm.counts.bg.4\_WT" "tpm.counts.bg.5\_WT" "tpm.counts.bg.6\_WT"

[57] "BS\_ID" "tpm\_support\_KO" "tpm\_support\_WT" "tpm\_supported"

🡪 Werte kommen aus differential binding

21-22: WT KO

🡪 ob die bindestelle in der jeweiligen condition vorhanden war (1 vorhanden, 0 nicht vorhanden)

23-25: "geneID.2" "geneName.1" "region.1"

🡪 die genannotation aus dem diff binding, falls sie anders ist als Spalte 9-12 besser Spalte 9-12 benutzen

26-38: counts.bs, counts.bg

🡪 number of crosslinks per sample that were found in the binding site or in the background

39-50: Deseq stats of differential binding

51-56 tpms of genes per sample

57 BS\_ID

->old binding site ID better not use

58-60 tpm support

🡪 additional differential binding stuff, stating wether the tpms are high enough for a differential analysis

61 down

🡪 whether binding site was downregulated in differential binding

62-69 seeds within 200nt downstream off binding site

62 table of all seeds 200nt downstream off binding site (position is relative to bs start)

63-66: "first\_seed\_200down.start" "first\_seed\_200down.end"

[65] "first\_seed\_200down.width" "first\_seed\_200down.type"

🡪 position of the downstream seed closest to binding site

67: "first\_seed\_200down.wobble"

Whether the first seed has the wobble position

68: "seed\_repetitions.200down" number of seeds in the 200nt window

69: "seed\_repetitions.200down.wobble" number of seeds in the 200nt window with a wobble position

70-75 same for seeds within 200nt upstream of binding site (position starts 200nt before the bs, so start 2 means starts -198)