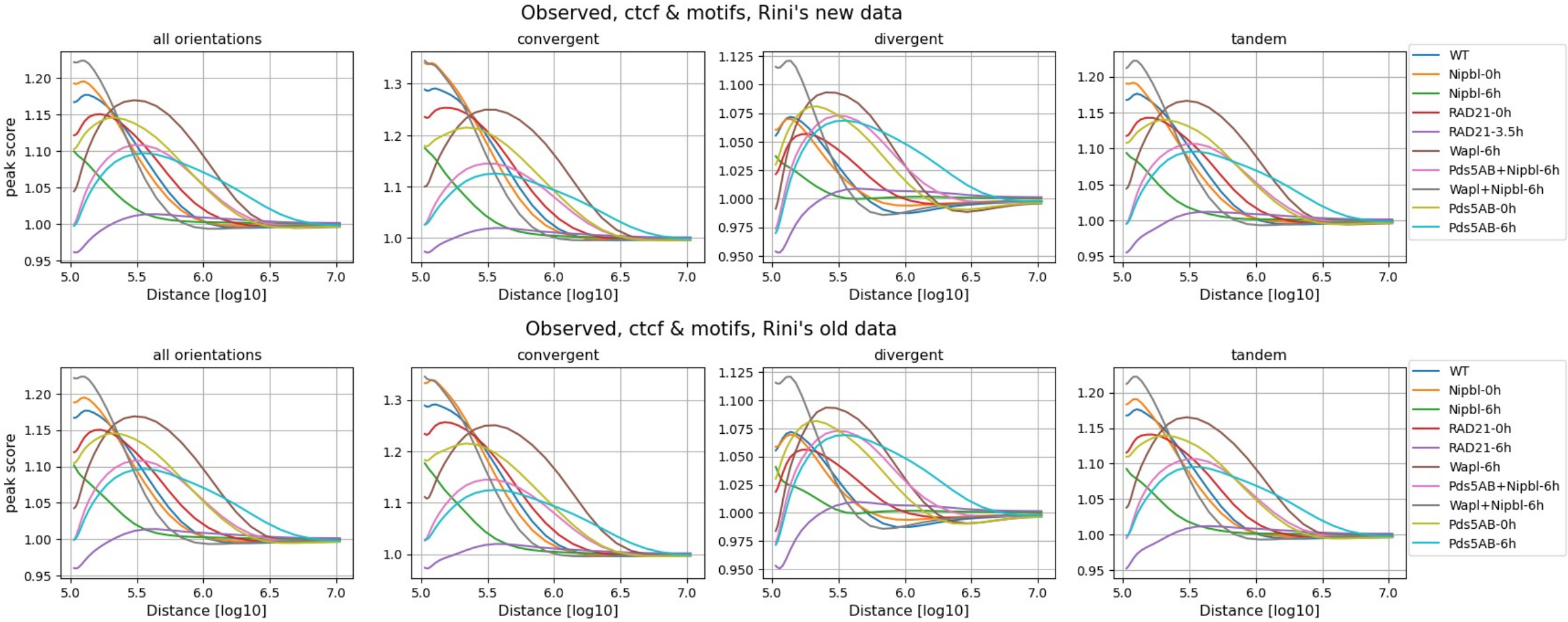
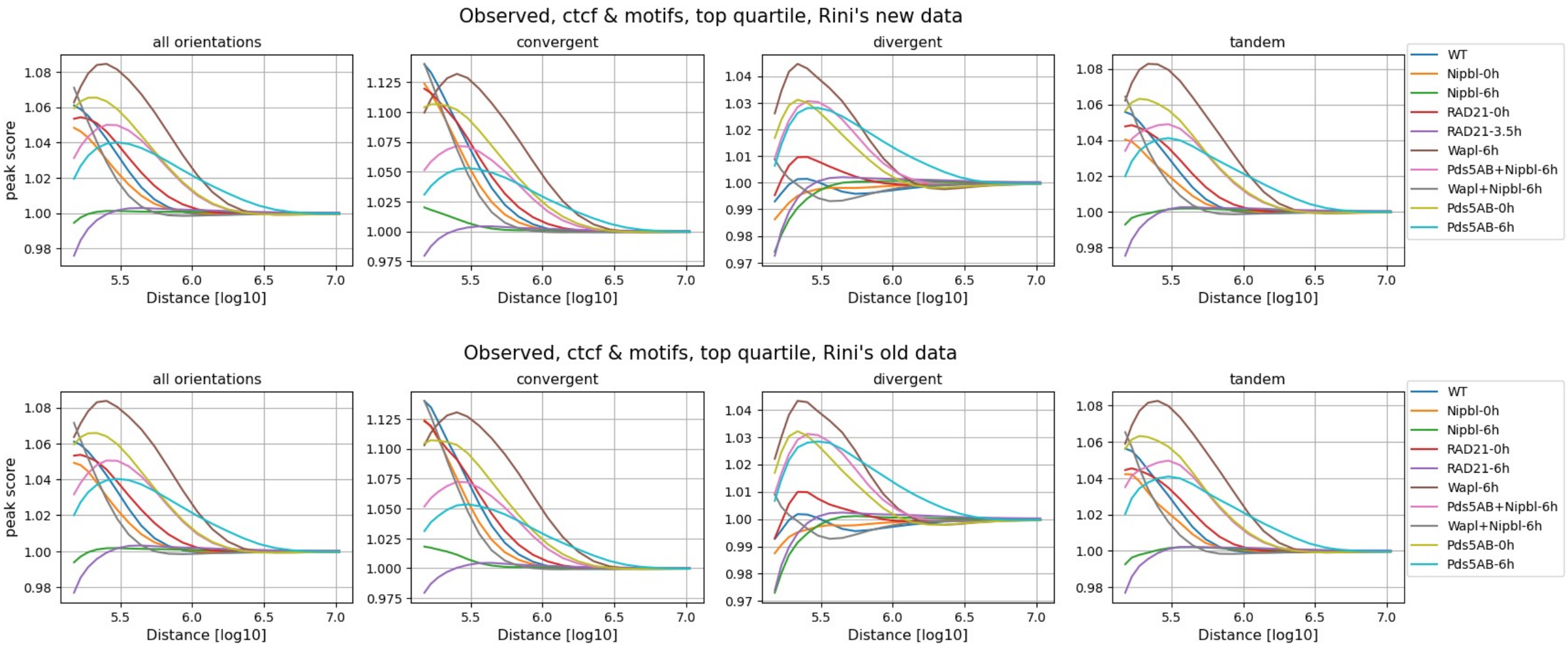
Peak analysis

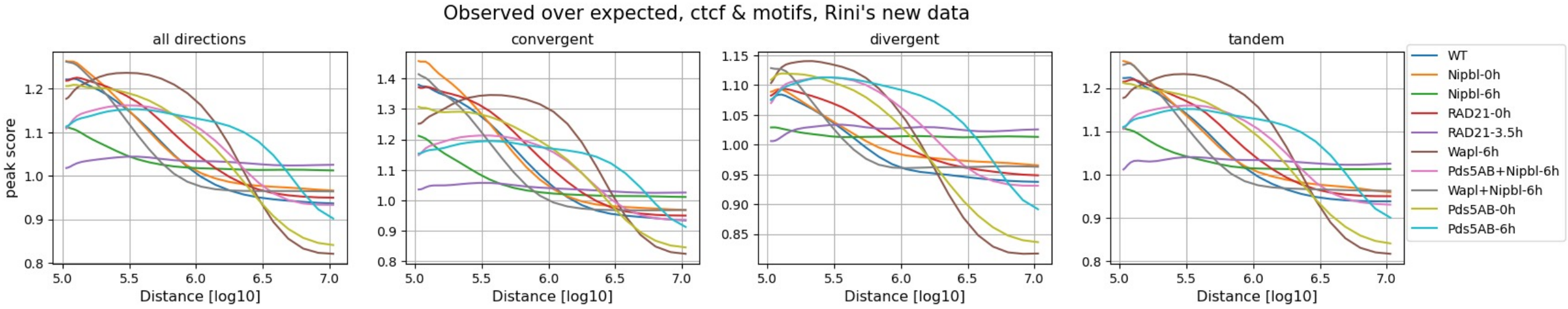
Jan 26, 2024

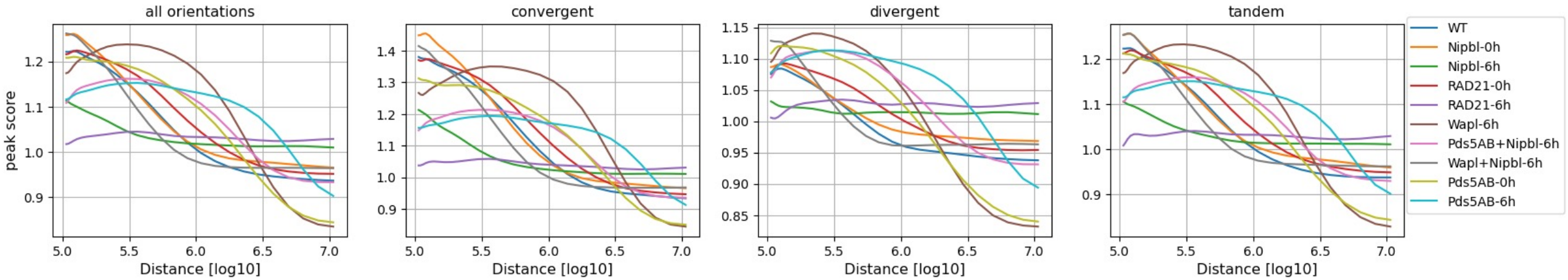


1. Plots are obtained by calculating the peak score constructed between any pair of positions where there is an overlap between CTCFs and motifs.
2. There are very similar trends between new data and old data; they are almost the same.
3. Here are figures for all orientations. While the interpretation might be complex, I think there might be information about the extruder number in divergence vs. convergence. For instance, in the convergent orientation, WT has a higher maximum magnitude compared to Wapl. However, in the divergent orientation, this happens inversely. The high score of WT in the convergent orientation can be due to the lower collision of extruders in its lower density.
4. I expected more similarity between Wapl and PDS5+Nipbl, but in fact, peaks are significantly stronger in Wapl. The magnitude of scores in Pds5+Nipbl is almost the same as Pds5 but with a lower extension. I think this might also be relevant to the number of extruders, as we observe different behavior in their divergent vs. convergent.
5. The same similarity might be expected between WT and Wapl + Nipbl, but again, in the divergent case, the difference is more considerable.
6. Such observations (3 and 4) might suggest that Wapl, and to a greater extent, PDS5, might weaken the extruders' anchor on chromatin.

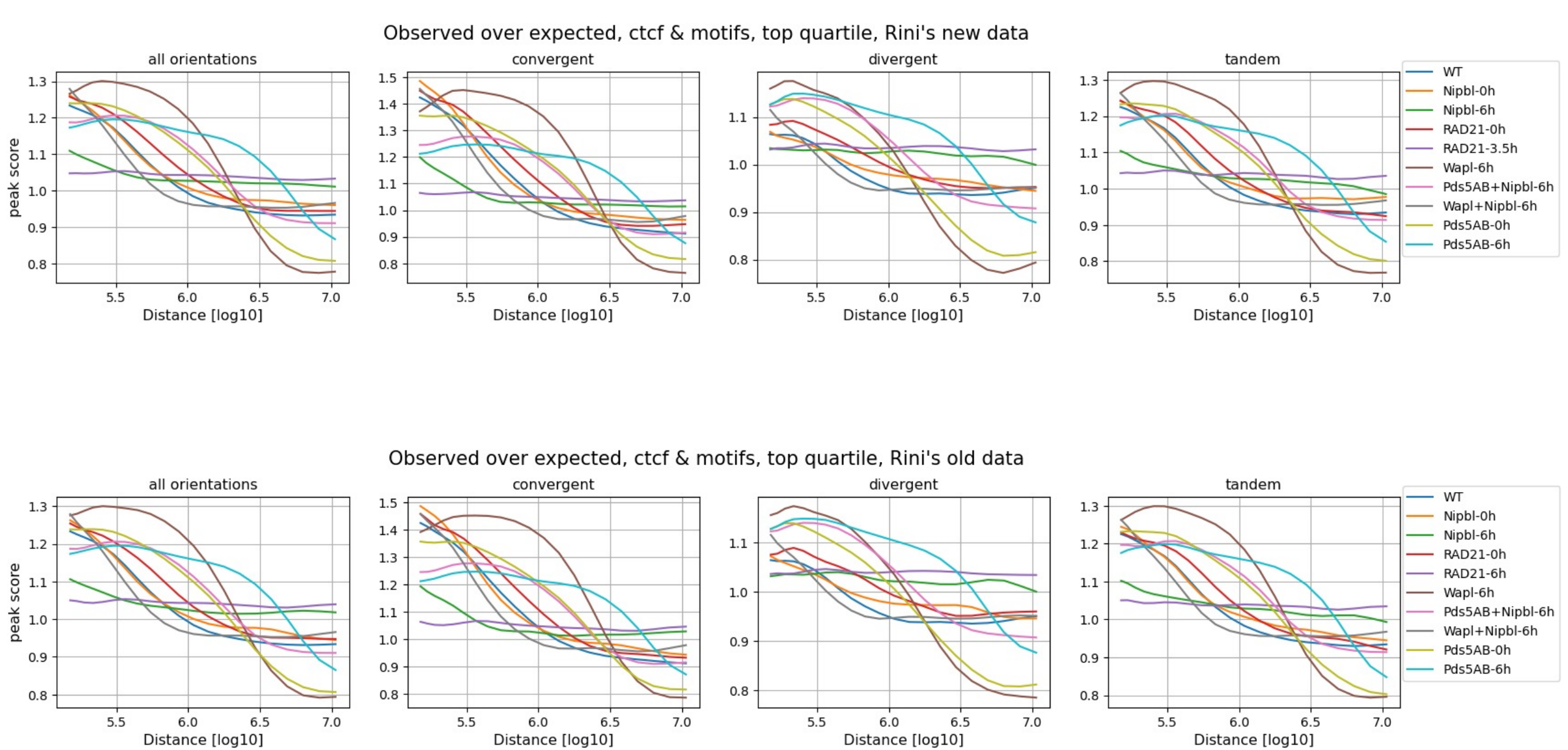


- The average distance between all CTCF (overlapping) sites is around 56 kb, but the average distance between CTCFs in the top quartile it is around 316 kb. Considering the place of the peak score maximum at 10^5.1~125 kb for WT, it means that if we only consider the top quartile motifs, we will not detect that.
- Such a larger distance, on average, leads to weaker scores.





- We expect the score to converge to one at a large distance. The lower score at such a range is due to the normalization procedure in "observed over expected" cases. The higher contacts of CTCF sites at peaks compensate by having lower contacts at larger distances, as the total number of contacts should be conserved.


Observed over expected, ctcf & motifs, top quartile, Rini's new data
all orientations
convergent
divergent
tandem
peak score
Distance [log10]
WT
Nipbl-0h
Nipbl-6h
RAD21-0h
RAD21-3.5h
Wapl-6h
Pds5AB+Nipbl-6h
Wapl+Nipbl-6h
Pds5AB-0h
Pds5AB-6h

Observed over expected, ctcf & motifs, top quartile, Rini's old data
all orientations
convergent
divergent
tandem
peak score
Distance [log10]
WT
Nipbl-0h
Nipbl-6h
RAD21-0h
RAD21-6h
Wapl-6h
Pds5AB+Nipbl-6h
Wapl+Nipbl-6h
Pds5AB-0h
Pds5AB-6h