

Cell Type	chr1	chr2	...	chr11	chr12	...	chr21	chr22	chrX	chrY
A549	Green	Green	...	Green		...				
GM12878	Green	Green	...	Green		...				
H1	Green	Green	...	Green		...				
HeLa			...		Purple	...	Purple	Purple	Purple	Purple
LNCaP	Green	Green	...	Green		...				
MCF7	Green	Green	...	Green		...				
THP1	Green	Green	...	Green		...				
HCT116	Blue	Blue	...	Blue		...				

Figure 2: **Example test scenario of our leave-one-out-training (LOOT) configuration.** This example that tests on OCRs from HeLa cells illustrates which chromosomes did we take OCRs from, for training (green), validation (blue) and testing (purple). In every scenario, OCRs from HCT116 cells from chromosomes chr1 thru chr11 were used for validation (the rest for this cell type were discarded), and we only trained on OCRs from those same chromosomes from all other cell types, testing only on OCRs from chromosomes chr12 thru chrY. This ensures the test set is truly novel every time, and prevents introducing any protocol or cell type-specific bias during training, as well as training on regions that could overlap in coordinates with those OCRs tested.