HiPiler: Visual Exploration of Large Genome Interaction Matrices with Interactive Small Multiples

Supplementary Material

Fritz Lekschas
lekschas@seas.harvard.edu
Harvard John A. Paulson School of Engineering and Applied Sciences

Benjamin Bach
bbach@seas.harvard.edu
Harvard John A. Paulson School of Engineering and Applied Sciences

Peter Kerpedjiev
pkerp@hms.harvard.edu
Harvard Medical School

Nils Gehlenborg
nils@hms.harvard.edu
Harvard Medical School

Hanspeter Pfister
pfister@seas.harvard.edu
Harvard John A. Paulson School of Engineering and Applied Sciences

IEEE Information Visualization (InfoVis), 2017

Supplementary Figures

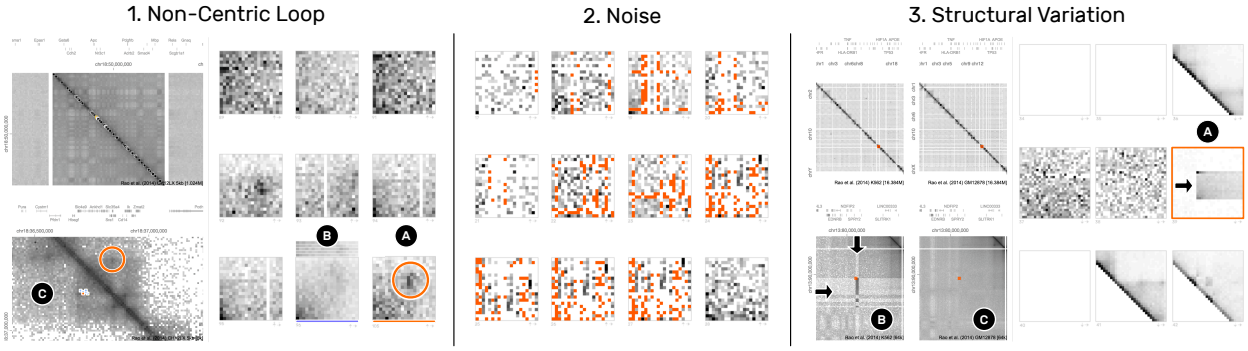

Figure S1: Three observations made during the data exploration sessions. (1) P1 and P3 individually found a group of pairwise enhancer-promoter interactions that are close to but not directly on a loop (1A and 1B). The loop pattern is indicated with an orange outline. 2A is the snippet with the most pronounced pattern. P1 and P3 wanted to remove the pile (2B) from the same location (C). (2) P4 investigated sparse snippets and visualized low quality cells (highlighted in orange) to identify that most of these snippets are extracted from a low quality region. (3) P5 highlighted a true positive structural DNA deletion (3A) by comparing two datasets (3B and 3C). The deletion (3B) causes the brighter (but not white) columns and rows (3B arrows) and is accompanied by an insertion, indicated by the dark rectangular area next to the highlighted location (3A arrow).

Supplementary Tables

Participant	Action	Task
P1	• Obtained overview through scrolling.	T1, T2
	• Ordered by snippets by size.	
	• Identified parts of the diagonal in snippets.	T1
	• Manually piled snippets via drag-and-drop.	T4
	• Manually piled snippets via swipe selection.	T4
	• Decreased size of snippets.	
	• Depiled all snippets.	
	• Arranged snippets with t-SNE.	T4
	• Manually piled noisy snippets via rectangular selection.	T4
	• Discarded noisy snippets.	
	• Identified area of snippets with a well-defined loop pattern.	T1
	• Manually piled snippets.	T4
	• Scaled up a pile of previously piled snippets to assess the aggregated signal.	T4
	• Leafed through the pile to ensure that it truly contains well-pronounced loop patterns.	T3
P2	• Obtained overview through scrolling.	T1
	• Arranged snippets by their distance to the diagonal and noise.	
	• Identified outliers, which appear to have no signal.	T2
	• Manually piled the outliers.	
	• Discarded the pile of outliers to increase the space for the other snippets.	
	• Identified noisy snippets, which are close to the diagonal.	T1
	• Automatically grouped all snippets that fall within the same grid cell.	
	• Manually piled the piles of noisy snippets close to the diagonal.	
	• Discarded the noisy pile.	
	• Increased size of all snippets to better see patterns.	T3
	• Scaled up and down individual snippets to inspect pattern.	T3, T4
	• Switched from arranging by noise to arranging by sharpness.	
	• Depiled all snippets and arranged them with t-SNE.	
	• Piled many noisy snippets using the swipe selection and discarded them.	
P3	• Obtained overview through scrolling.	
	• Identified loop-like pattern.	T1, T2
	• Clicked on the snippet to highlight the location in the matrix view.	
	• Opened detail matrix view and zoomed into the snippet location.	
	• Identified non-centric loop-like pattern.	T2, T3
	• Tried to correlate to external metadata. [This data was not available in the user study.]	T5
	• Found several co-located snippets exhibiting similar patterns.	T1
	• Manually piled up all but one of the co-located snippets and discarded them.	
	• Arranged snippets with t-SNE.	
	• Identified large area of snippets with a noisy pattern.	T2
	• Identified area of snippets with well-defined interaction patterns.	T1, T2

	• Manually piled up snippets with the lasso tool.	T4
	• Inspected snippets on the pile up.	T3
P4	• Obtained overview through scrolling.	
	• Identified that many snippets are sparse.	T1
	• Activated the visualization of low-quality bins to distinguish between sparsity and low quality.	T3
	• Wanted to group by number of low-quality bins.	
	• Arranged snippets by size.	
	• Arranged snippets with t-SNE.	
	• Identified group of co-located snippets.	
	• Piled up all co-located snippets previously identified.	T4
	• Located the pile of snippets in the matrix view and confirmed that all snippets are related to another pattern.	T1, T3
P5	• Obtained overview through scrolling.	T1, T2
	• Enabled visualization of low quality bins.	
	• Manually piled up snippets of low quality using the swipe and rectangular tool	
	• Discarded pile of low quality snippets.	
	• Reverted discarding and piling as they wanted to inspect snippets with no patterns.	T1, T2
	• Opened detail matrix view.	
	• Inspected one snippet in the detail matrix.	T3
	• Loaded second data set and navigated to the same region in the snippet view from before to compare the ROI.	T6
	• Identified snippet as a false positive pattern after inspecting the context of both snippets.	T3
	• Inspected a second snippet region in both dataset's detail matrix views.	T3, T6
	• Identified the region as a true positive DNA deletion event.	T1

Table S1: Chronological summary of participant-specific actions and related tasks of the user study.