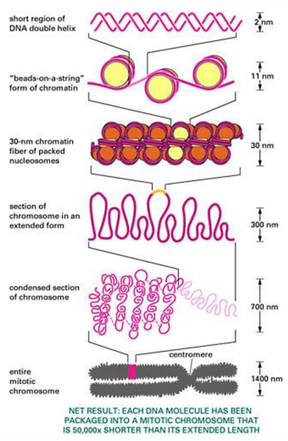
Genome-wide 3D maps of regulatory interactions in the mouse developing forebrain

Shira Ginzburg[1], Noga Moshe[2], Chew Yee Ngan[3,4], Elkana Baris[1], Yubo Zhang[3,4], Reut Eshel[2], Chia-Lin Wei[3,4], Nadav Ahituv[5], Ramon Birnbaum[2] and Tommy Kaplan[1]

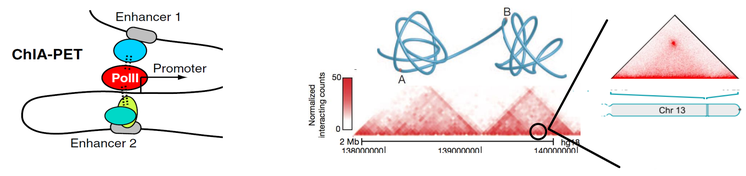
[1]The Hebrew University, [2]Ben-Gurion University, [3]Berkeley Lab, [4]DOE Joint Genome Institute, [5]University of California, San Francisco.

1 Studying the role of genome packaging in gene regulation



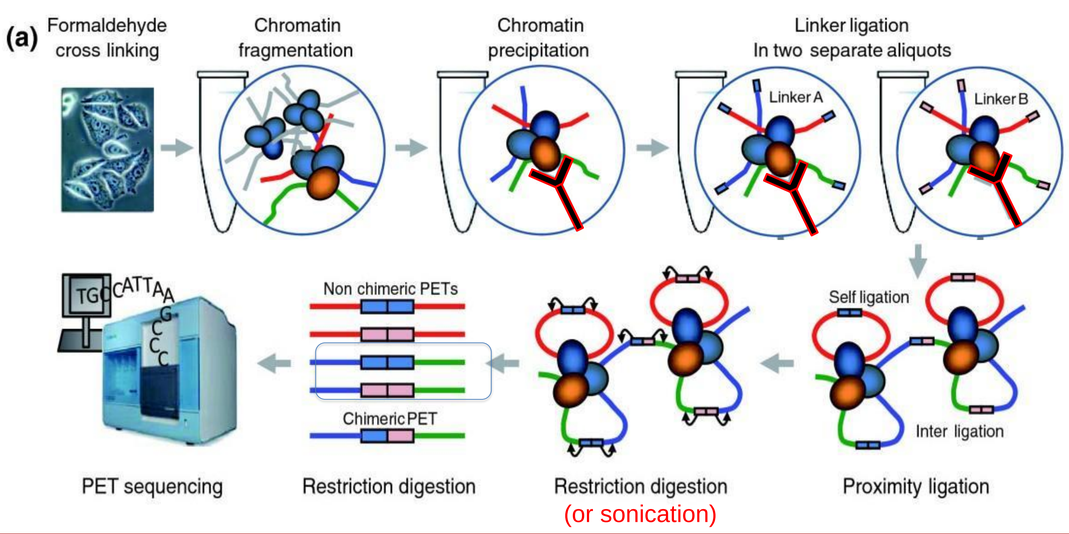
Existing methods for mapping 3D genome interactions:

- FISH - limited in throughput and resolution.
- 3C-based methods (including Hi-C) - **regulatory functional interactions** are often obscured among many non-functional DNA-DNA interactions.



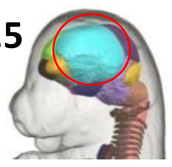
2 ChIA-PET

- ChIA-PET - experimental way of measuring genome-wide DNA-DNA interaction mediated by protein of interest.

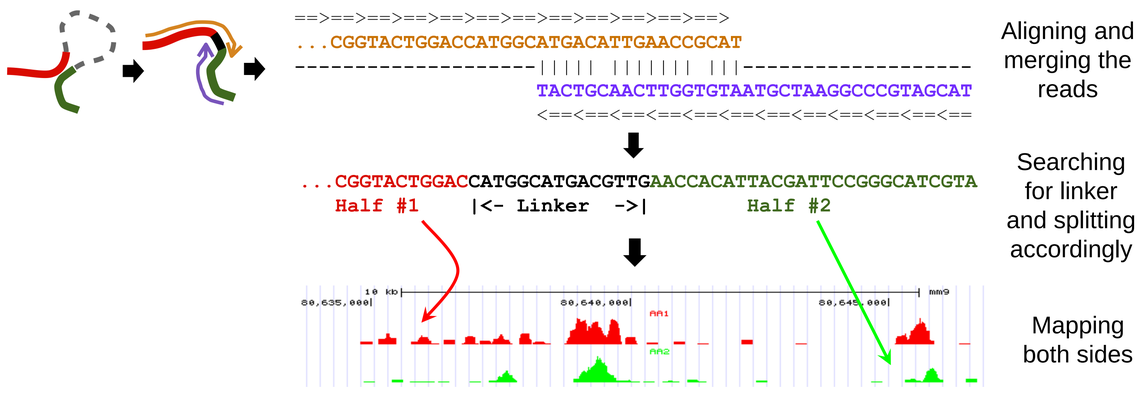


3 Data parsing and mapping

- **ChIA-PET of RNA PolII and H3K27ac in mice brain E16.5**
- **Goal: Genome-wide identification of regulatory interactions in the mouse developing forebrain**



- **Input**: 498 million 150bp paired-end reads.

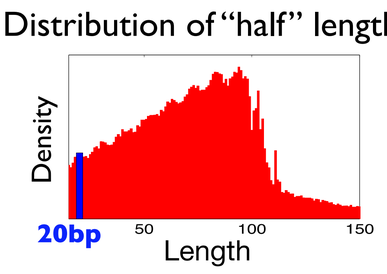


	H3K27ac			PolII		
	AA	BB	AB	AA	BB	AB
Number of sequenced pairs	257M			241M		
Number of overlapping pairs (≥20bp)	192M (75%)			206M (80%)		
Mapped (up to 10 places)	21M	20M	13.6M	9.7M	9.1M	3.9M

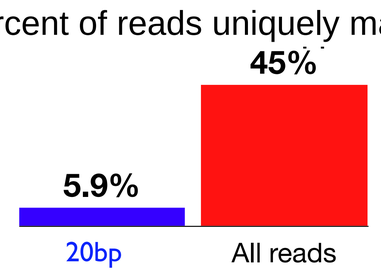
- Mapping: Remove reads with more than 10 alignments. From best alignments, choose pair on same chr / min dist

- **Comparison to standard MmeI (20bp long) ChIA-PET**

Distribution of "half" lengths



Percent of reads uniquely mapped



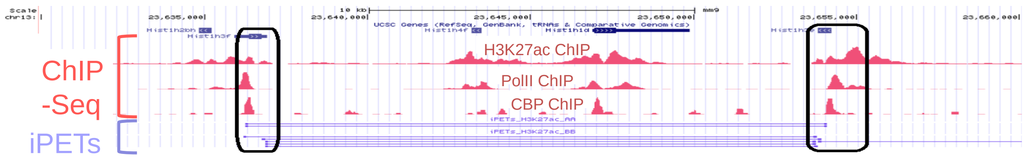
4 Obtaining a high-confidence set of interactions

(1) Filter out:

- Interactions between different chromosomes
- Self-ligation interactions (dist<1200 bp and same strand)
- Long-range interactions (dist>1Mbp)

	H3K27ac			PolII		
	AA	BB	AB	AA	BB	AB
Mapped	21M	20M	13.6M	9.7M	9.1M	3.9M
Same chr	4.7M	4.8M	912K	4.8M	5.0M	500K
[Dist>1.2K | diff. strand]	1M	967K	780K	388K	343K	224K
Dist<1M	42K	42K	15K	78K	78K	7.8K
Overlaps ChIP peaks (=iPETs)	12K	11K	4K	35K	35K	2K
Percent of sequenced	0.009 %			0.03 %		

(2) Use the in vivo binding landscape of PolII, CBP, and the H3K27ac to prune our set of interactions:

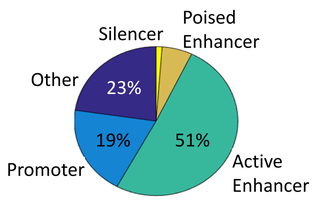


Peak set: peaks from ChIP-Seq experiments of **PolII**, **H3K27ac** and **CBP** from mice brain E16.5 (~90,000 peaks)

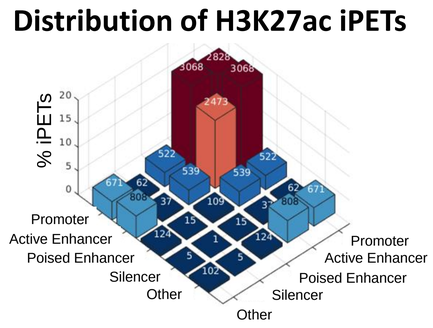
Retain only **iPETs = same-chromosome peak-to-peak interactions**.

5 Interactions enriched for **enhancer-promoter** and **promoter-promoter** interactions

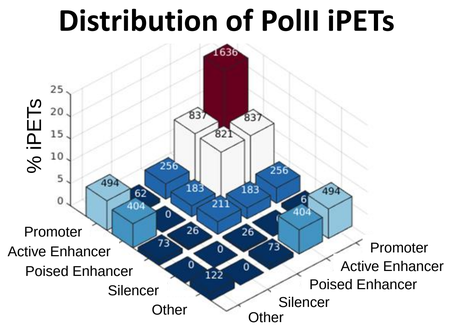
Peak Annotations



Distribution of H3K27ac iPETs



Distribution of PolII iPETs

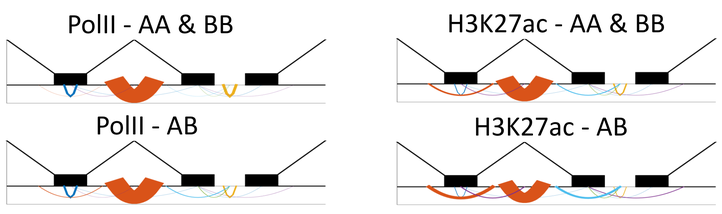


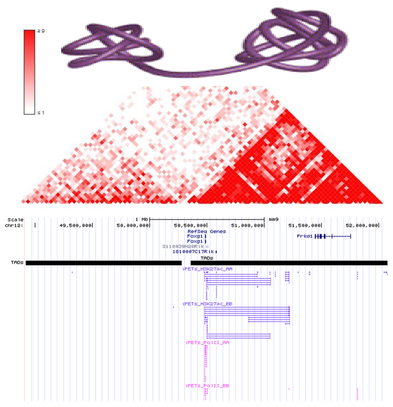
- **Promoter**: -2.5 Kb +500 bp w.r.t TSS
- **Active Enhancer**: H3K27ac (and not H3K27me3)
- **Poised Enhancer**: H3K27ac and H3K27me3
- **Silencer**: H3K27me3 (and not H3K27ac)

In total, we found in our data ~1,000 promoters which interact (#iPETs≥2) with regulatory regions (enhancers/silencers)

6 Most interactions are within TADs (94%)

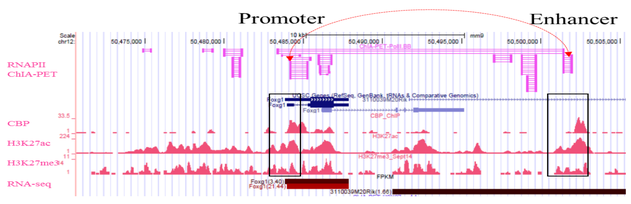
Distribution of iPET types



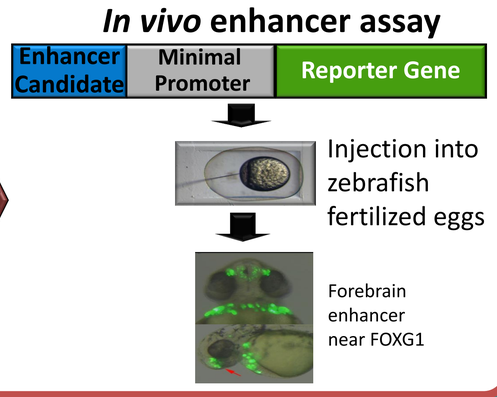


7 *In vivo* enhancer assay

Identifying Genome-Wide Enhancer Candidates



In vivo enhancer assay



8 References

1. Lieberman-Aiden *et al*, *Science*, 2009
2. Dixon *et al*, *Nature* 2012
3. Shen *et al*, *Nature* 2012
4. Fullwood *et al*, *Nature* 2009
5. Fullwood *et al*, *Gen. Res.*, 2009
6. Zhang *et al*, *Nature*, 2013