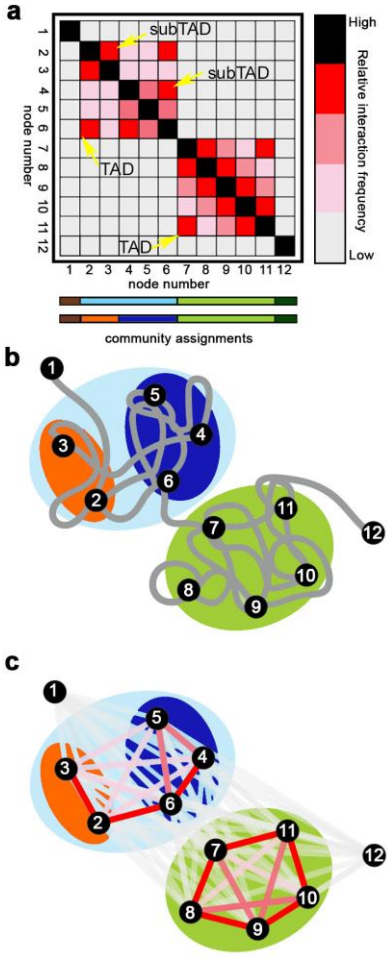
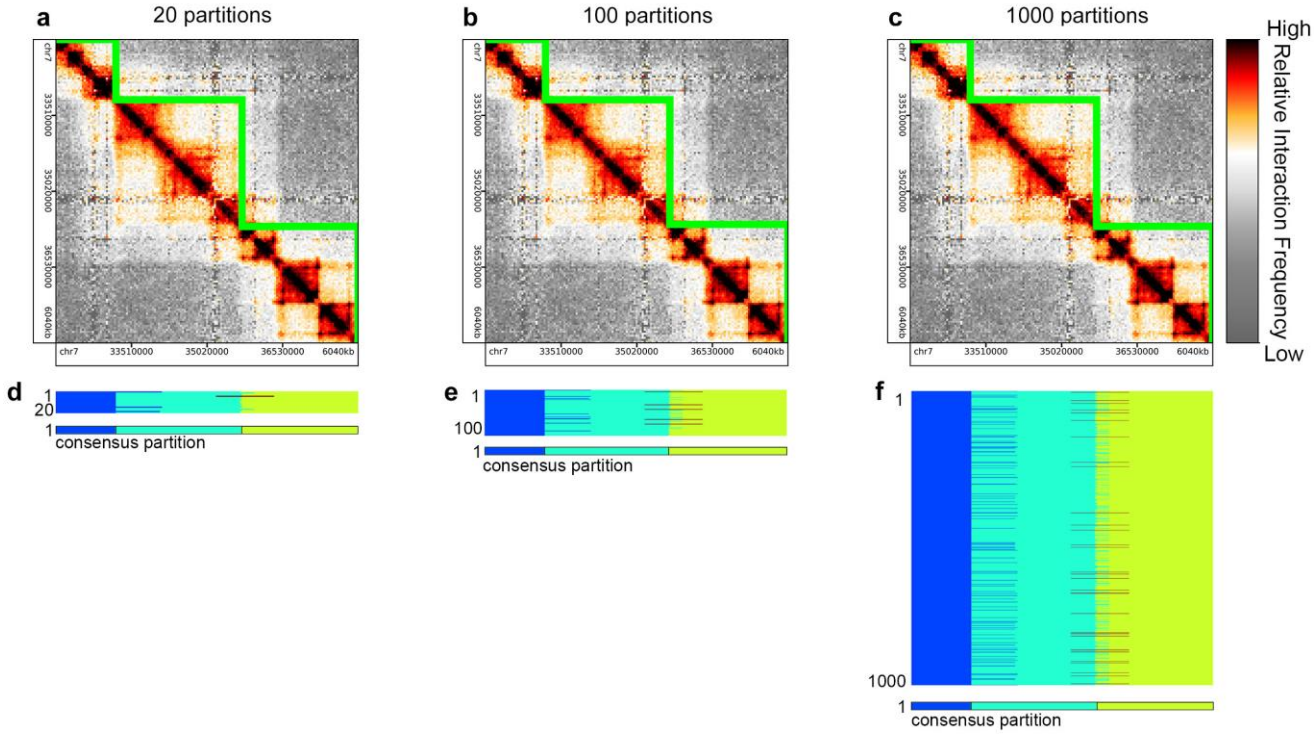


Supplementary Figure 1

3D genome folding data can be represented mathematically as a network.

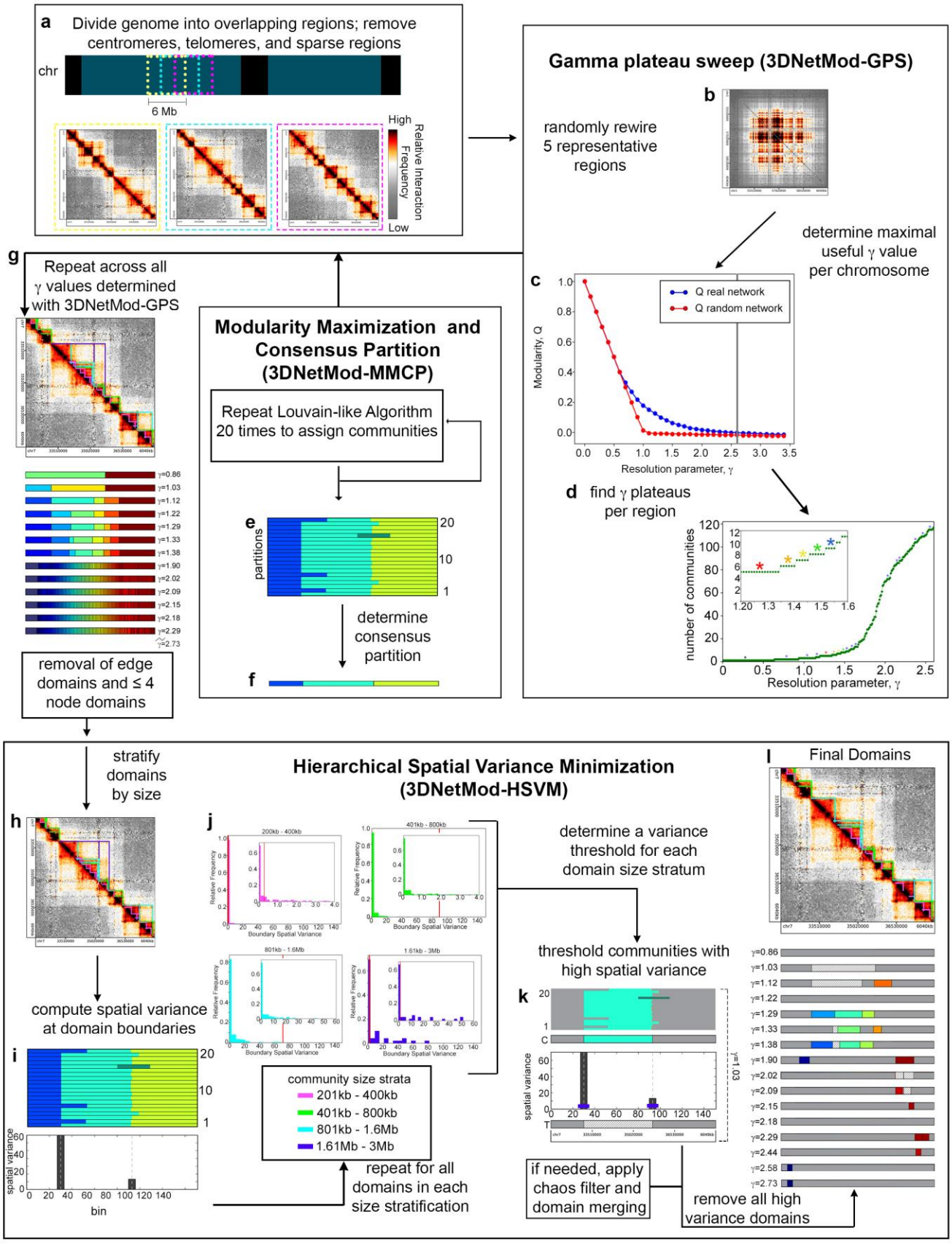
(**A**) Schematic representation of a Hi-C heatmap with Megabase-scale topologically associated domains (TADs) and nested sub-TADs within the first TAD. Node numbers correspond to genomic bins. Yellow arrows indicate the corners of domains. (**B**) Model of the potential community structure underlying proximity ligation data depicted in panel **(A)**. Black dots with numbers represent nodes corresponding to the genomic coordinates represented in each bin. (**C**) Spring force diagram representation of a genome-folding network. Edge colors depict edge weight ranging from low (grey) to high (red).



Supplementary Figure 2

Effect of partition number in 3DNetMod-MMCP on consensus partition.

(A-C) An approximately 6 Mb Hi-C region of human cortical plate cells[1]. Communities identified using 3DNetMod-MMCP with **(A)** 20 underlying partitions, **(B)** 100 underlying partitions and (**C**) 1000 underlying partitions at a γ value of 1.03 are outlined in green. (**D-F**) The underlying individual partitions in the partition block are used to determine the consensus partition with a γvalue of 1.03.

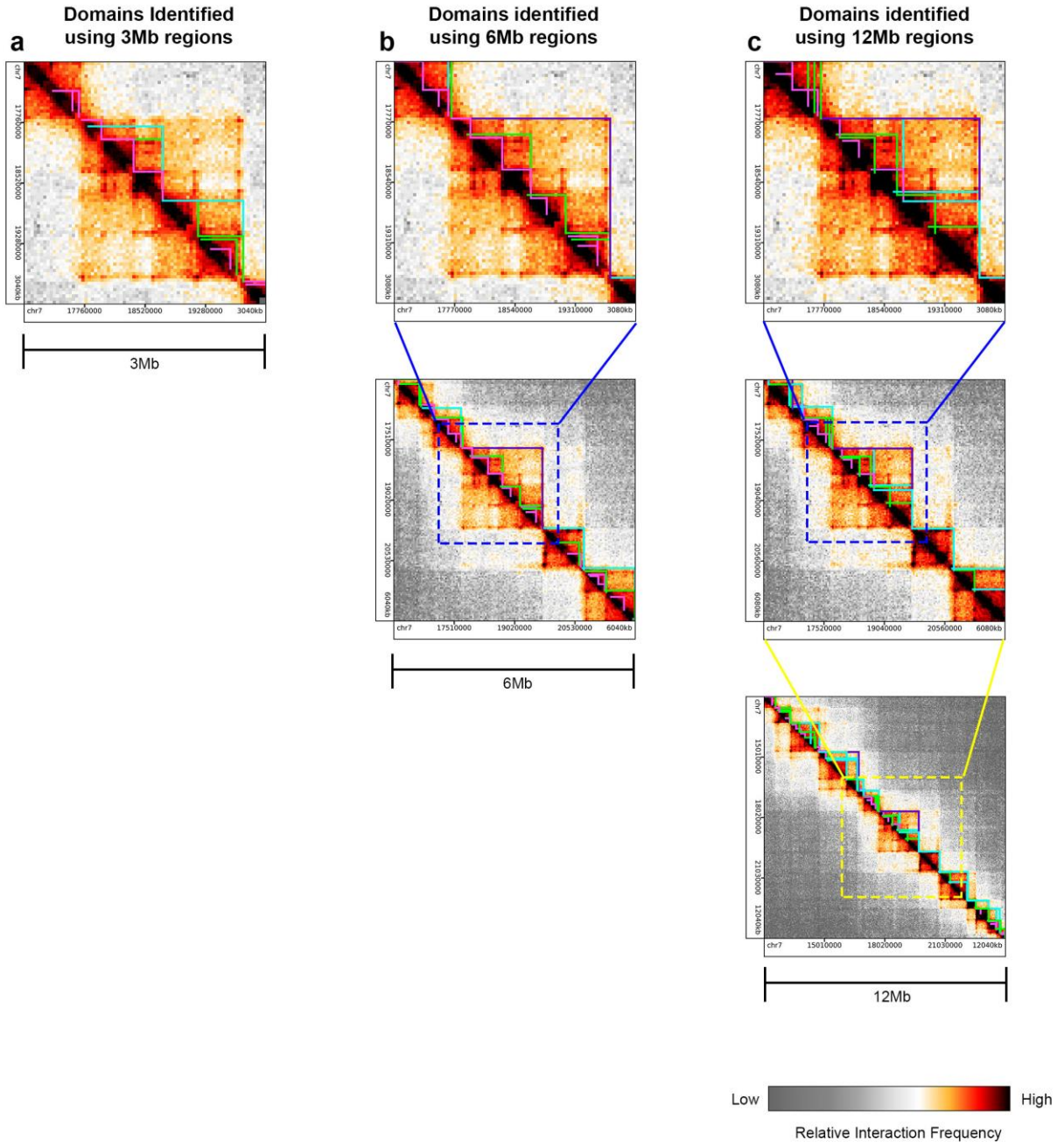


Nature Methods: doi:10.1038/nmeth.4560

Supplementary Figure 3

Overview of 3DNetMod method.

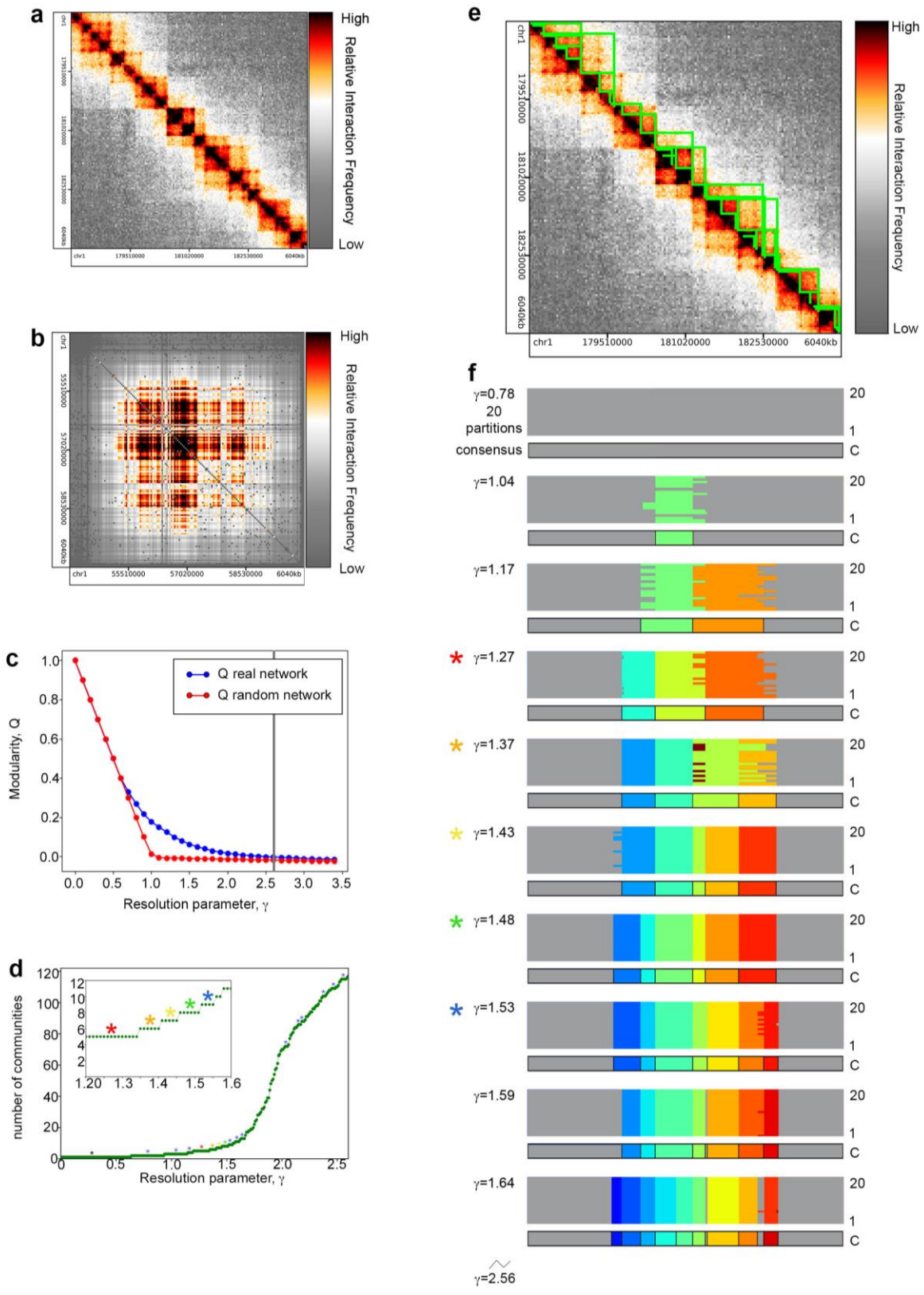
(**A**) Each chromosome is parsed into overlapping regions with user-defined size and overlap. (**B-D**) Gamma Plateau Sweep (3DNetMod-GPS). **(B)** Five representative regions equally distributed along the chromosome are randomly rewired to create null networks. **(C)** Modularity is computed across a range of γ values for the five representative regions and their randomly rewired counterparts. The convergence point between the modularity curves of real and random networks is determined for each region (grey vertical line). The maximal informative γ value is computed for each chromosome as the maximum γ at the convergence points for the five regions. **(D)** To identify γ values that give rise to high-confidence communities, we defined 'plateaus' of finely incremented (0.01 step size) γ values that give rise to the same number of communities. The midpoints for all plateaus that are greater than a user-defined minimal plateau size (indicated with a star) are selected as γ values for high-confidence community detection. (**E-F**) Modularity maximization and consensus partition (3DNetMod-MMCP). **(E)** For each selected γ value, the network is partitioned 20 times into communities. **(F)** A consensus partition is computed from the 20 partitions. **(G)** 3DNetMod-MMCP is repeated across all γ values selected with 3DNetMod-GPS, resulting in the initial detected set of communities across all length scales. Small communities (less than or equal to 4 nodes) and communities at region edges are removed. **(H-L)** Hierarchical Spatial Variance Minimization (3DNetMod-HSVM). **(H)** First, communities are stratified by size: ≤400kb (magenta), 401kb – 800kb (green), 801kb – 1600kb (cyan), 1601kb – 3Mb (indigo) and > 3Mb (orange). **(I)** Spatial variance is then computed at each community boundary. **(J)** Boundary spatial variance values for each community of a given size stratification across the entire chromosome are pooled. A user-defined variance threshold is computed for each size-stratified set of communities. **(K)** If a community has a boundary with higher spatial variance than the size-specific threshold, it is removed from further consideration. **(L**) The resulting variance-thresholded communities can be further refined by merging significantly overlapping domains or domains called in regions devoid of nested substructure (see Supplemental Methods). The final community list represents high-confidence 3DNetMod domains.



Supplementary Figure 4

Effect of region size on domains detected.

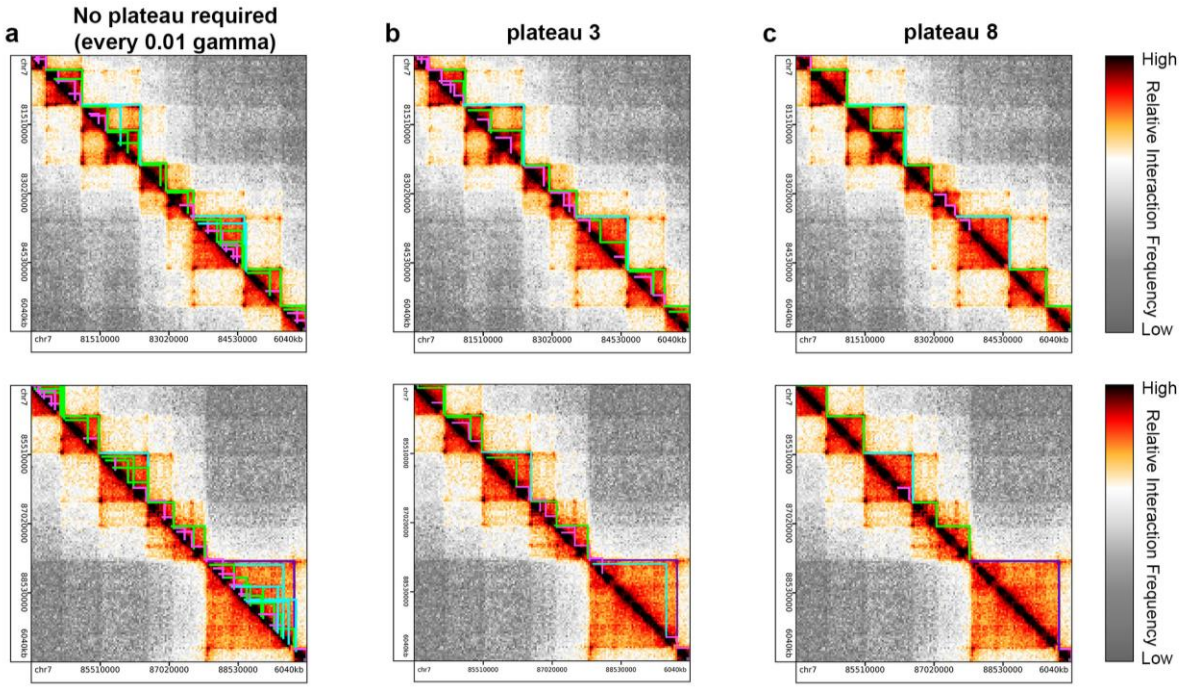
(**A**) Representative 3 Mb Hi-C heatmap from human cortical plate[1]. Domains identified with 3DNetMod-GPS, 3DNetMod-MMCP and 3DNetMod-HSVM using 3Mb regions are outlined. (**B**) Top panel: Heatmap of the same 3Mb in (**A**) with domains identified using 6Mb outlined. Bottom panel: The full 6Mb region used for region detection. (**C**) Domains identified using 12Mb regions.



Supplementary Figure 5

Gamma plateau sweep (3DNetMod-GPS) sensitively partitions Hi-C data into a hierarchy of partially overlapping TADs and subTADs genome-wide.

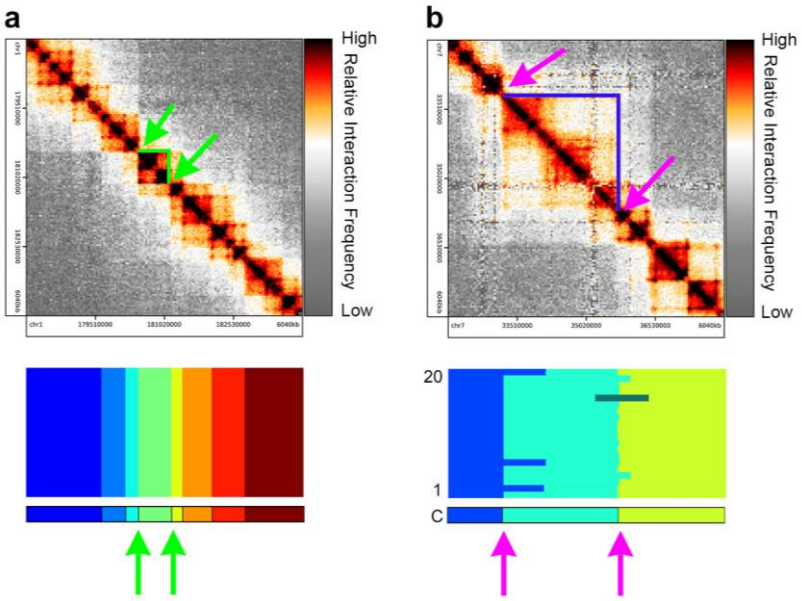
(A) A 6 Mb-sized Hi-C region from human cortical plate tissue[1]. **(B)** Representative randomly rewired 6Mb Hi-C network in which the locations of existing edges have been reassigned according to expected edge weights. (**C**) Modularity, *Q, versus* structural resolution parameter, γ, for a representative 6Mb region and its corresponding randomly re-wired network. The maximal useful γ value is the point of convergence between real and random networks and is indicated by a vertical grey line. **(D)** Number of communities as a function of γ for the Hi-C region shown in **(A)**. Three or more consecutive γ values with the same number of communities are grouped into plateaus. The midpoint of each plateau (indicated by a star) is selected as a valid γ for community detection. **(E)** Heatmap with domains identified across the selected resolution parameter values. **(F)** Partition blocks (partitions 1 through 20) and consensus partition ('C') for community assignments across the selected γ values. Domains that contact the edge of the network and domains < 200kb have been removed from consideration and are greyed out on partition blocks.



Supplementary Figure 6

Effect of plateau method on 3DNetMod domain detection.

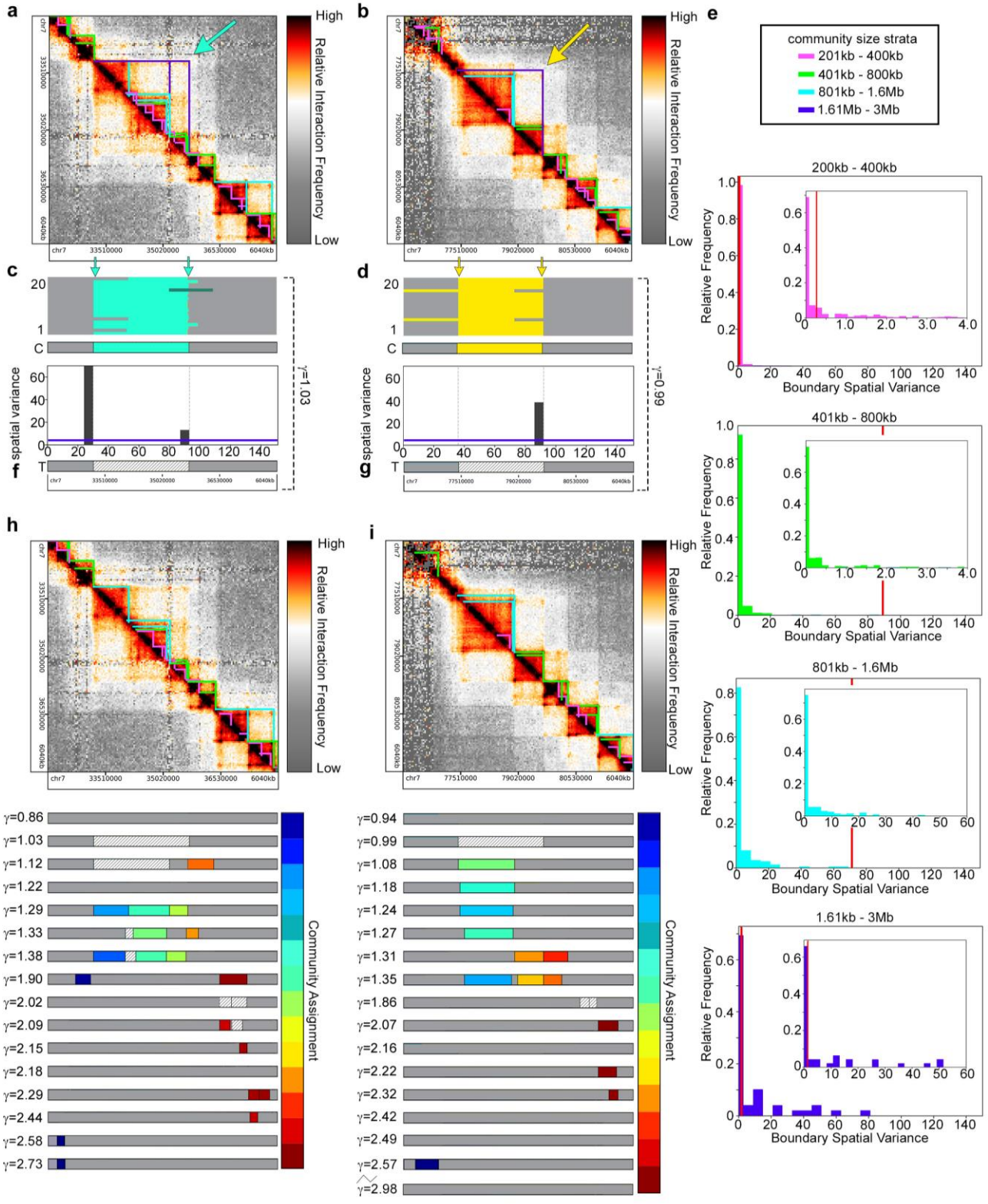
(A-C) Hi-C heatmaps of 2 representative 6 Mb regions (2 rows) from human embryonic cortical plate tissue[1] with communities detected by 3DNetMod-GPS and 3DNetMod-MMCP **(A)** without the plateau method or with plateau sizes of **(B)** 3 or **(C)** 8.



Supplementary Figure 7

High variance at potential A/B compartment boundaries.

(A) A 6 Mb Hi-C region from human embryonic cortical plate tissue[1]. The blocks representing 20 partitions and the corresponding consensus partition are shown beneath the heatmap. Community assignments have no variability across the 20 partitions. An illustrative domain with zero variability in assignment across the 20 partitions is indicated with green arrows. (**B**) Another 6 Mb Hi-C region from human embryonic cortical plate tissue[1]. The boundaries of a putative A/B compartment-like structure exhibit high spatial variance across 20 partitions (magenta arrows).

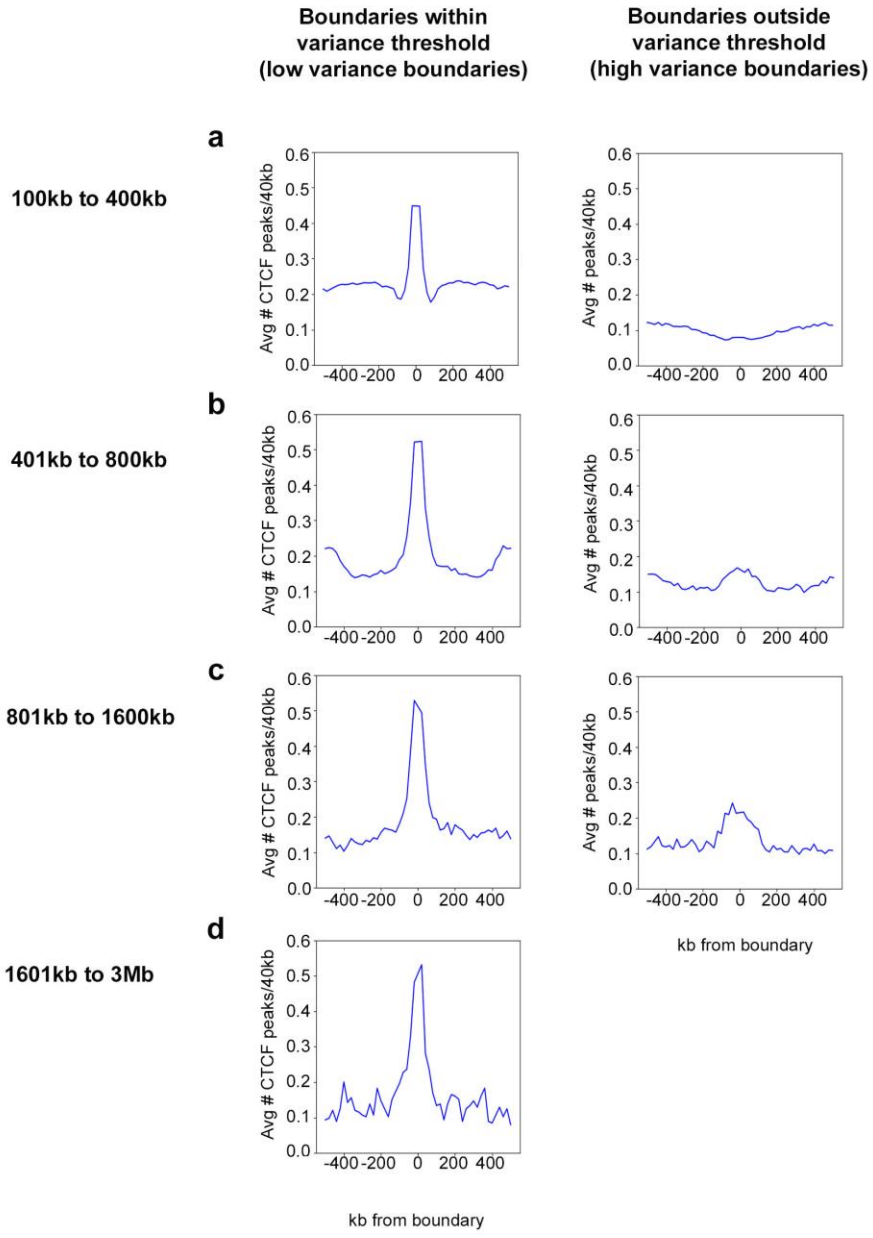


Supplementary Figure 8

Hierarchical spatial variance minimization (3DNetMod-HSVM) dissects *bona fide* chromatin domains from higher-order architectural

features.

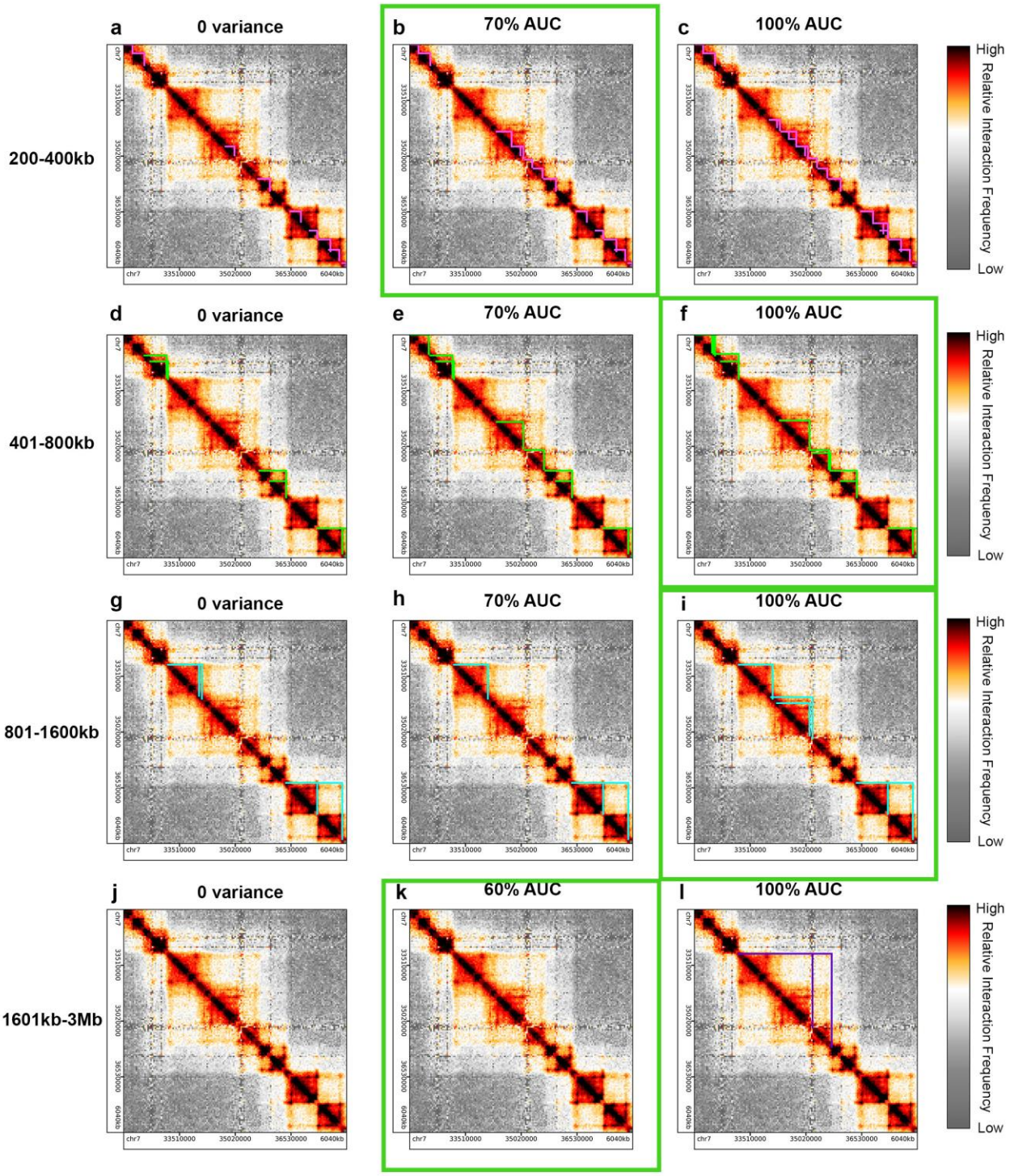
(A-B) Two different 6Mb genomic regions from human cortical plate tissue Hi-C[1] after gamma plateau sweep (3DNetMod-GPS) to sensitively partition a hierarchy of partially overlapping TADs and subTADs genome-wide. A putative A/B compartment with alternating, non-uniform signal is indicated with an arrow in each heatmap. **(C-D)** Twenty partitions and consensus partition (labeled as 'C') for γ values at which the compartments illustrated in A-B, respectively, were detected. Domain calls at edges of the network are removed and colored in grey in the partition block and consensus partition. The spatial variance across the 20 partitions for each boundary is computed. **(E)** Histograms of spatial variance from all domain boundaries from chromosome 7 stratified by the size of domain**.** A variance threshold can be selected for each domain size stratum (red vertical line) to filter higher-order compartments and low-confidence domains from the full detected set. **(F-G)** Domain calls are removed from consideration (thresholded consensus partition, 'T') if either of its boundaries exceeds the length scale-specific variance threshold. **(H-I)** Final nested hierarchy of domains after 3DNetMod-HSVM.



Supplementary Figure 9

CTCF enrichment at high and low variance boundaries.

(**A-D**) Average number of wild type mouse neuronal CTCF peaks[2] per 40kb centered at boundaries of domains identified with 3DNetMod. Boundaries of domains in the **(A)** 100 - 400 kb, **(B)** 401 - 800 kb and **(C)** 801 -1600 kb size stratum with low (left) or high variance and were removed with variance thresholding (right). (**D**) Boundaries of domains in the 1601 kb - 3 Mb size stratum. No variance thresholding was performed at this size stratum.

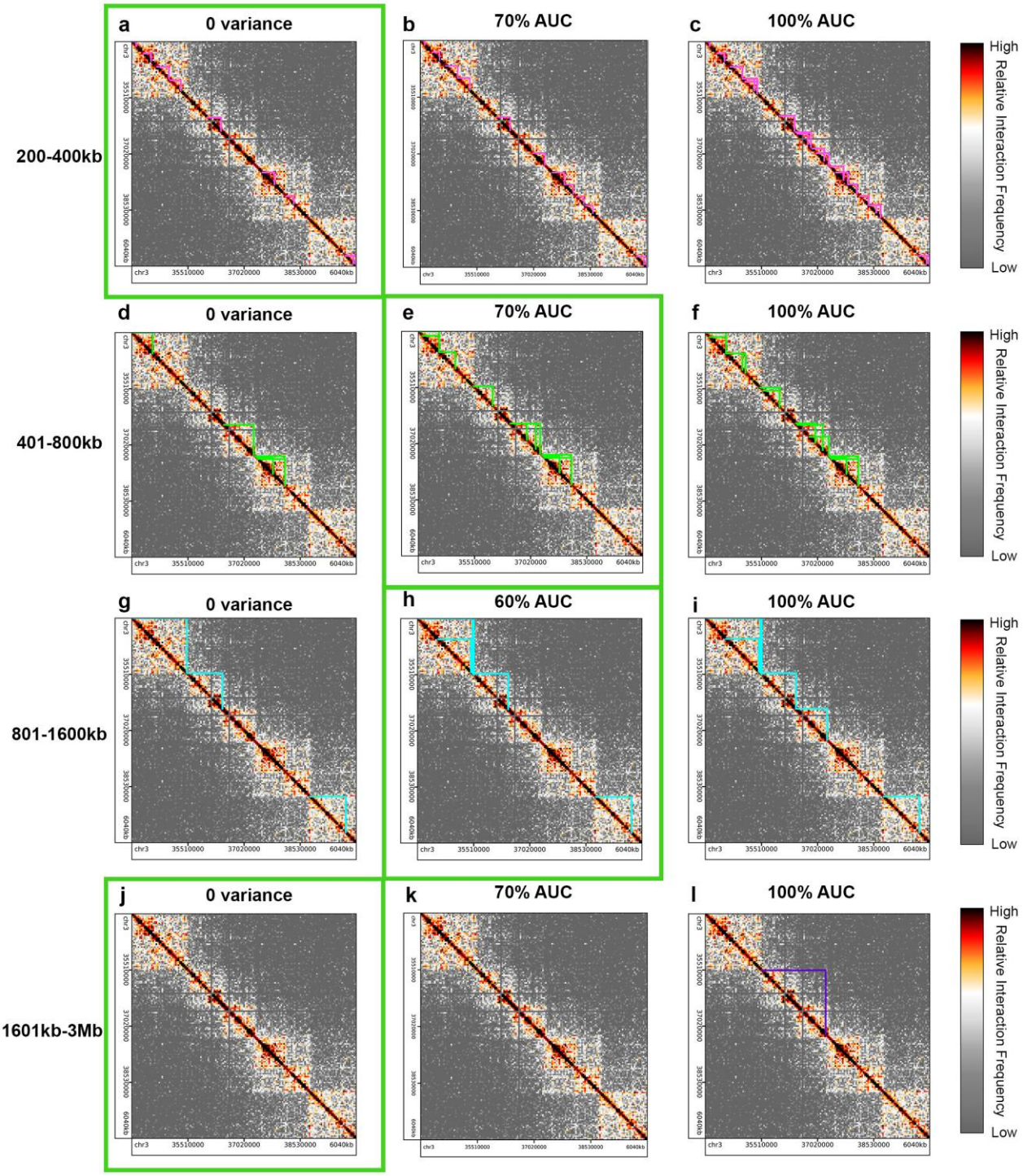


Supplementary Figure 10

Selection of variance thresholds on human embryonic cortical plate tissue Hi-C[1].

(**A-C**) 6 Mb Hi-C heatmaps from human embryonic cortical plate tissue binned at 40 kb[1]. Domains identified with 3DNetMod-GPS and

3DNetMod-MMCP at the 200 – 400 kb size stratum with 3 different variance thresholds during hierarchical spatial variance minimization (3DNetMod-HSVM). Domains at (**A**) 0 variance, (**B**) 70% AUC and (**C**) 100% AUC thresholds outlined in magenta. The selected threshold is indicated with a green box. (**D-L**) 3 different variance thresholds at the (**D-F**) 401 – 800 kb size stratum, (**G-I**) 801 -1600 kb and (**J-L**) 1601 kb – 3 Mb size strata.

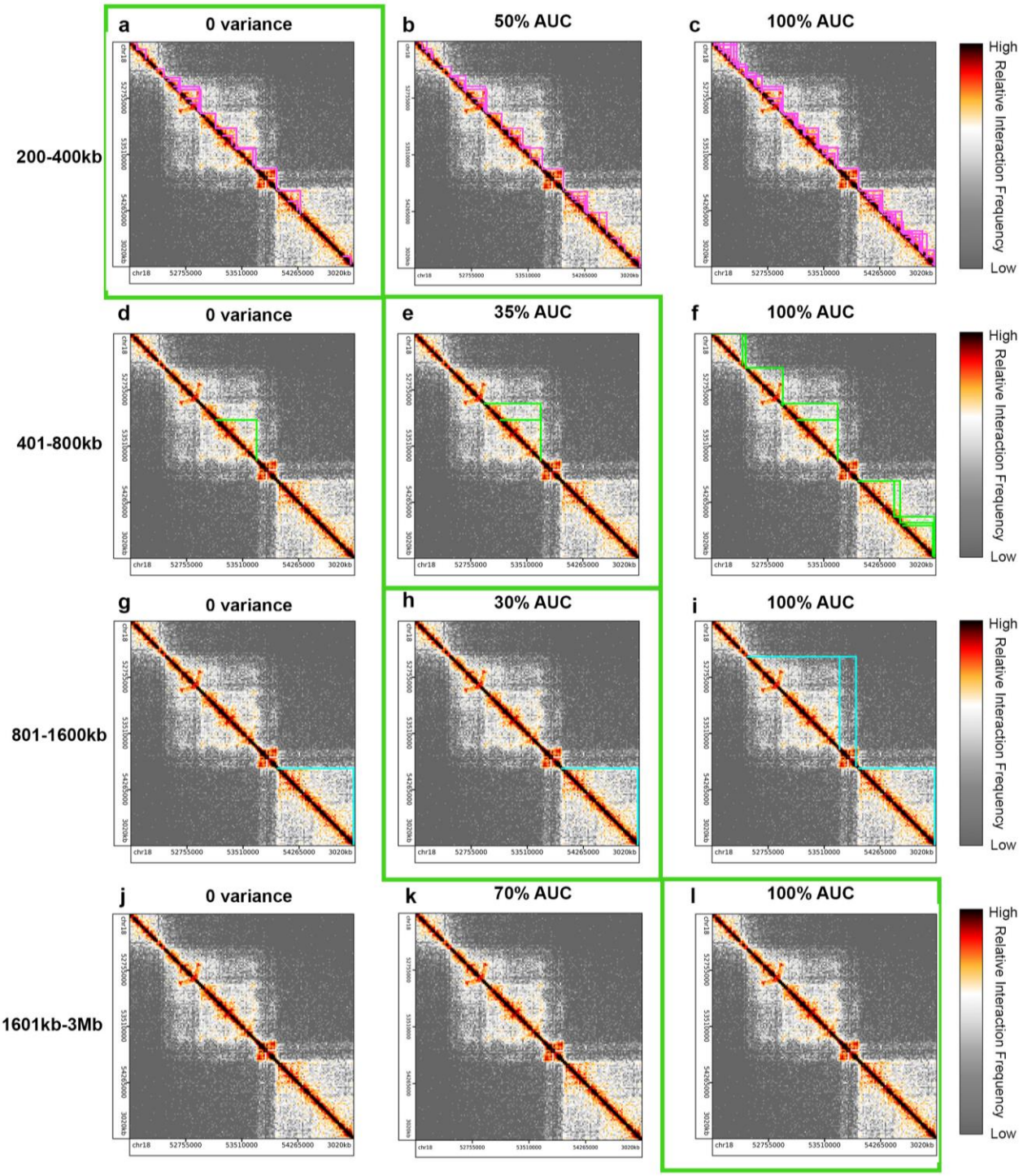


Supplementary Figure 11

Selection of variance thresholds on mouse cortical tissue Hi-C[3].

(**A-C**) 6 Mb Hi-C heatmaps from mouse cortical tissue binned at 40 kb[3]. Domains identified with 3DNetMod-GPS and 3DNetMod-

MMCP at the 200 – 400 kb size stratum with 3 different variance thresholds during hierarchical spatial variance minimization (3DNetMod-HSVM). Domains at (**A**) 0 variance, (**B**) 70% AUC and (**C**) 100% AUC thresholds outlined in magenta. The selected threshold is indicated with a green box. (**D-L**) 3 different variance thresholds at the (**D-F**) 401 – 800 kb, (**G-I**) 801-1600 kb and (**J-L**) 1601 kb – 3 Mb size strata.

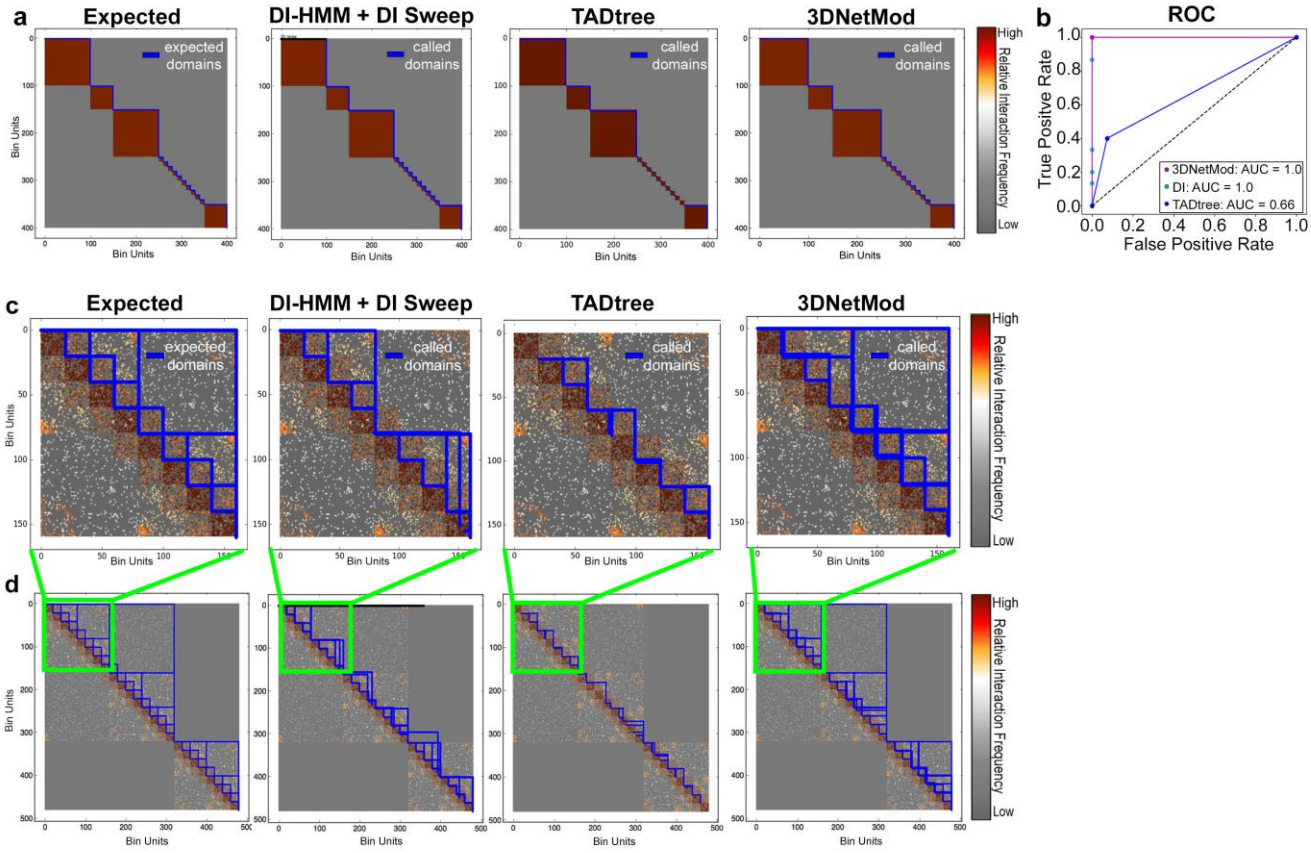


Supplementary Figure 12

Selection of variance thresholds on mouse neuron Hi-C[2].

(**A-C**) 3 Mb Hi-C heatmaps from mouse neurons binned at 20 kb[2]. Domains identified with 3DNetMod-GPS and 3DNetMod-MMCP at the 200 – 400 kb size stratum with 3 different variance thresholds during hierarchical spatial variance minimization (3DNetMod-HSVM). Domains at (**A**) 0 variance, (**B**) 70% AUC and (**C**) 100% AUC thresholds outlined in magenta. The selected threshold is indicated with a

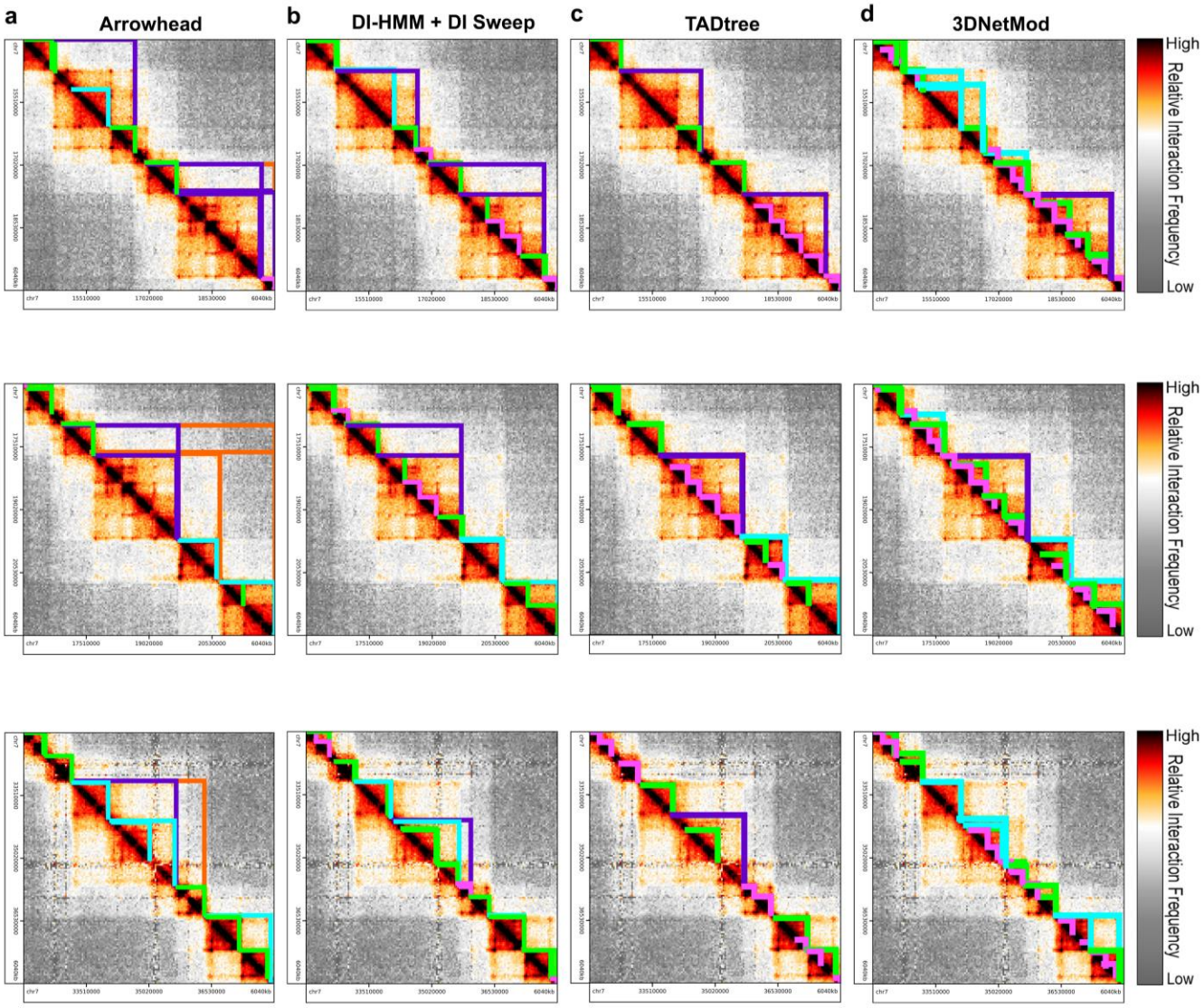
green box. (**D-L**) 3 different variance thresholds at the (**D-F**) 401 – 800 kb, (**G-I**) 801-1600 kb and (**J-L**) 1601 kb - 3Mb size strata.



Supplementary Figure 13

Comparison of domain-calling performance across methods on simple simulated Hi-C data (corresponding to Figure 2).

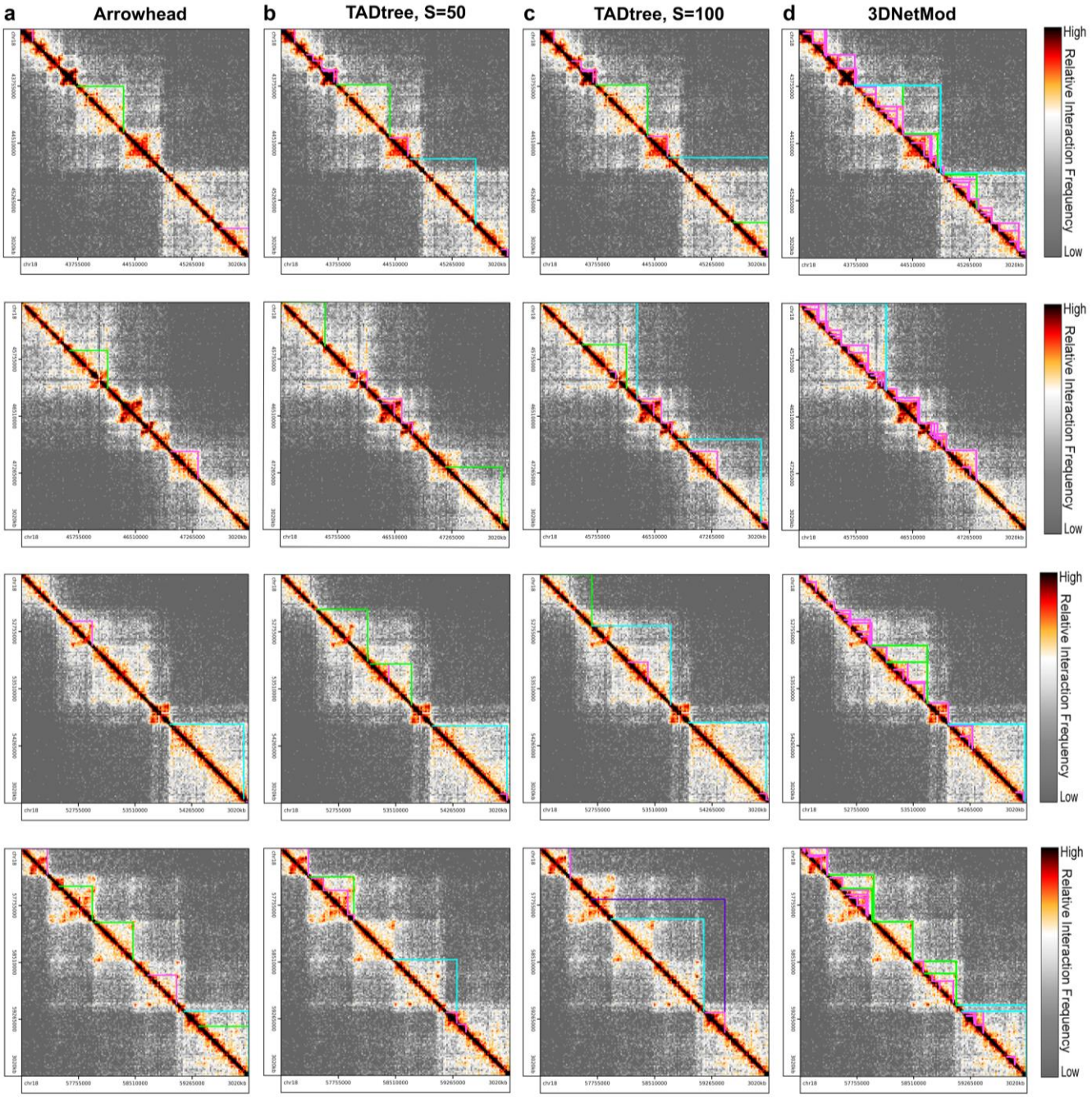
(A) Simulated Hi-C data with simple, non-overlapping domains of variable size. Expected domains (left), DI-HMM + DI sweep (center left), TADtree (center right) and 3DNetMod (right) are shown. (**B**) Receiver Operating Characteristic (ROC) curves showing the true positive rate and false positive rate of 3DNetMod (magenta), DI-HMM + DI Sweep (teal) and TADtree (blue) domain detection performance in the simple simulated Hi-C network. (**C**) Simulated Hi-C data with nested, partially overlapping domain structure. Expected domains (left), DI-HMM + DI sweep domains (center left), TADtree domains (center right) and 3DNetMod domains (right) are shown. (**D**) Full view of the zoomed in simulations shown in (**C**) (green box).



Supplementary Figure 14

Comparison of domains identified in 40 kb binned human embryonic cortical plate tissue Hi-C[1] by leading 3D chromatin domain detection methods.

Three representative 6 Mb Hi-C heatmaps from human embryonic cortical plate tissue[1]. Domains outlined as identified by **(A)** Arrowhead[4], (**B**) DI-HMM Domain Caller[3] with a sweep of directionality index (DI) values, (**C**) TADtree[5] and (**D**) 3DNetMod.



Supplementary Figure 15

Comparison of domains identified in 20 kb binned mouse neuron Hi-C[2] by leading 3D chromatin domain detection methods.

(A-D) Four representative 3 Mb Hi-C heatmaps from mouse neurons[2]. Domains outlined as identified by **(A)** Arrowhead[4], (**B**) TADtree[5] with S=50, (**C**) TADtree with S=100 default parameter and (**D**) 3DNetMod. Due to intractable run times of TADtree on 20kb-bnined data with S=100, TADtree was run on half of chromosome 18.

Tables

Supplementary Table 1. Publicly Available Hi-C Data Re-analyzed in Norton, Emerson et al.

Sample	Reference	GEO ID / Link
Human Cortical plate	Won et al. 2016 doi:10.1038/nature19847	GSM2054564 GSM2054565 GSM2054566 ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE77nnn/GSE77565/suppl/GSE77565_FBD_IC-heatmap-chr-40k.hdf5.gz
Human Germinal zone	Won et al. 2016 doi:10.1038/nature19847	GSM2054567 GSM2054568 GSM2054569 ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE77nnn/GSE77565/suppl/GSE77565_FBP_IC-heatmap-chr-40k.hdf5.gz
Mouse Embryonic Stem, Rep1	Dixon et al. 2012 doi: 10.1038/nature11082	GSM862720
Mouse Embryonic Stem, Rep2	Dixon et al. 2012 doi: 10.1038/nature11082	GSM862721
Mouse Cortex, Rep1	Dixon et al. 2012 doi: 10.1038/nature11082	GSM892304
Mouse Cortex, Rep2	Dixon et al. 2012 doi: 10.1038/nature11082	GSM892305
Mouse Wild-type neuronal, Rep1	Jiang et al. 2017 doi: 10.1038/ng.3906	GSM2643068
Mouse Wild-type neuronal, Rep1	Jiang et al. 2017 doi: 10.1038/ng.3906	GSM2643070
Mouse Setdb1 deficient neuronal, Rep1	Jiang et al. 2017 doi: 10.1038/ng.3906	GSM2643067
Mouse Setdb1 deficient neuronal, Rep2	Jiang et al. 2017 doi: 10.1038/ng.3906	GSM2643071

Supplementary Table 2: 3DNetmod time scaling comparison to TADtree**

	40 kb matrix resolution		20 kb matrix resolution	
	chr 7	**Genome-wide**	**chr 18**	**Genome-wide**
TADtree	2.21 hrs	44.5 hrs*	69.44 hrs*	> 63 days*
3DNetmod	0.86 hrs	22.4 hrs	0.87 hrs	1.82 days

* Estimate based on single chromosome.
**Time in terms of 4 processors in quad core (2.5 GHz Intel Core i7), default settings

Supplementary Table 3. Publicly available ChIP-seq data re-analyzed in Norton, Emerson et al.

Target	Cell type	Mapped Test ChIP-Seq reads after down-sampling	reference	Test Sample GEO ID	Control Samples	Mapped Control ChIP-seq reads after down-sampling	Control Sample GEO ID
CTCF	WT mouse neuronal	42E6	Jiang et al. 2017[20] doi: 10.1038/ng.3906	GSM2643058	WT mouse neuronal Whole Cell Extract	61E6	GSM2643066