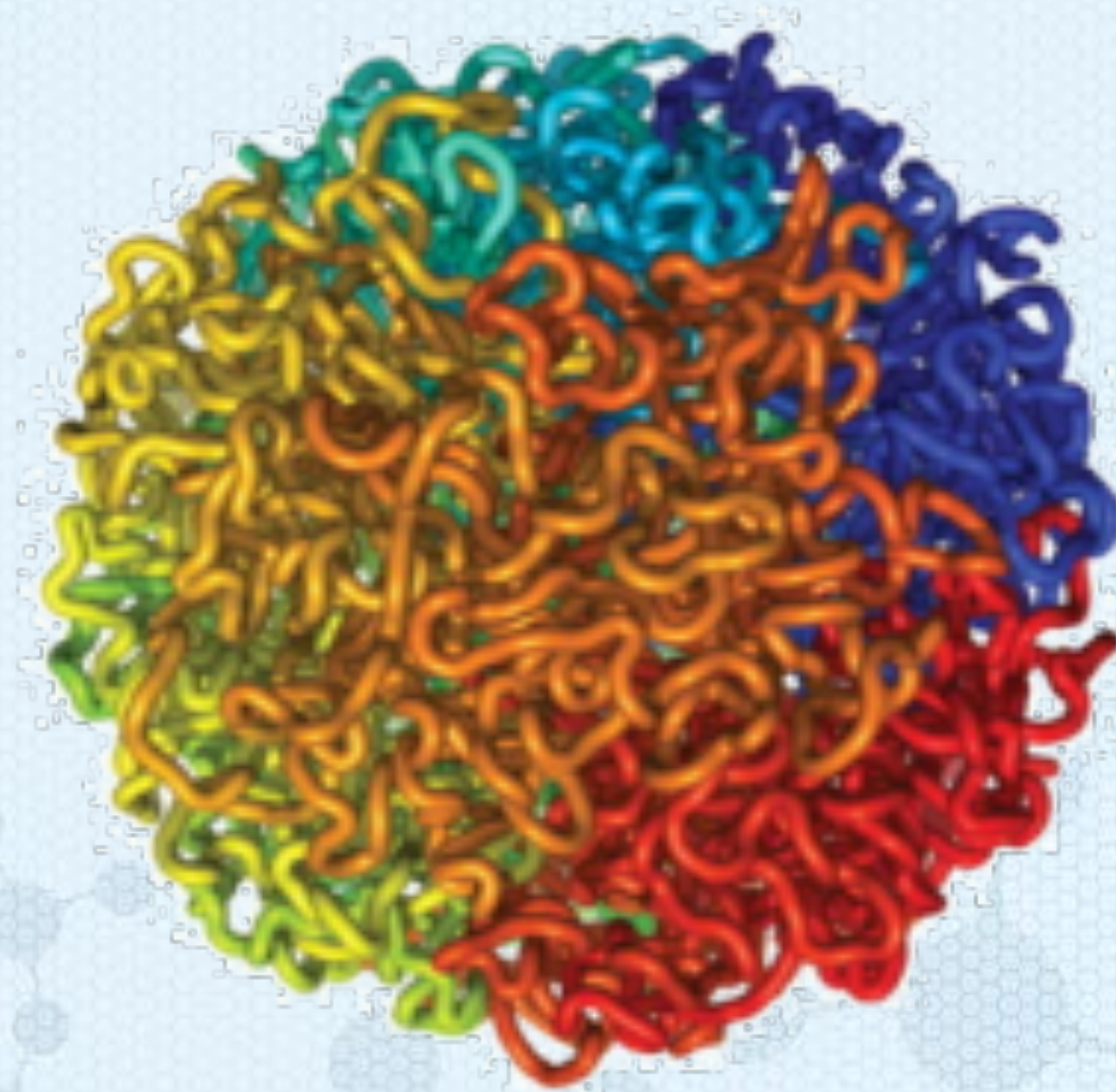


Building high quality genomes with PacBio + HiC



@PhaseGenomics

@Zevkronenberg

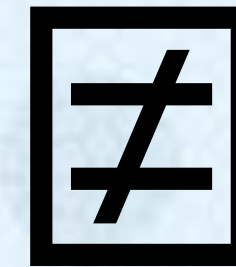
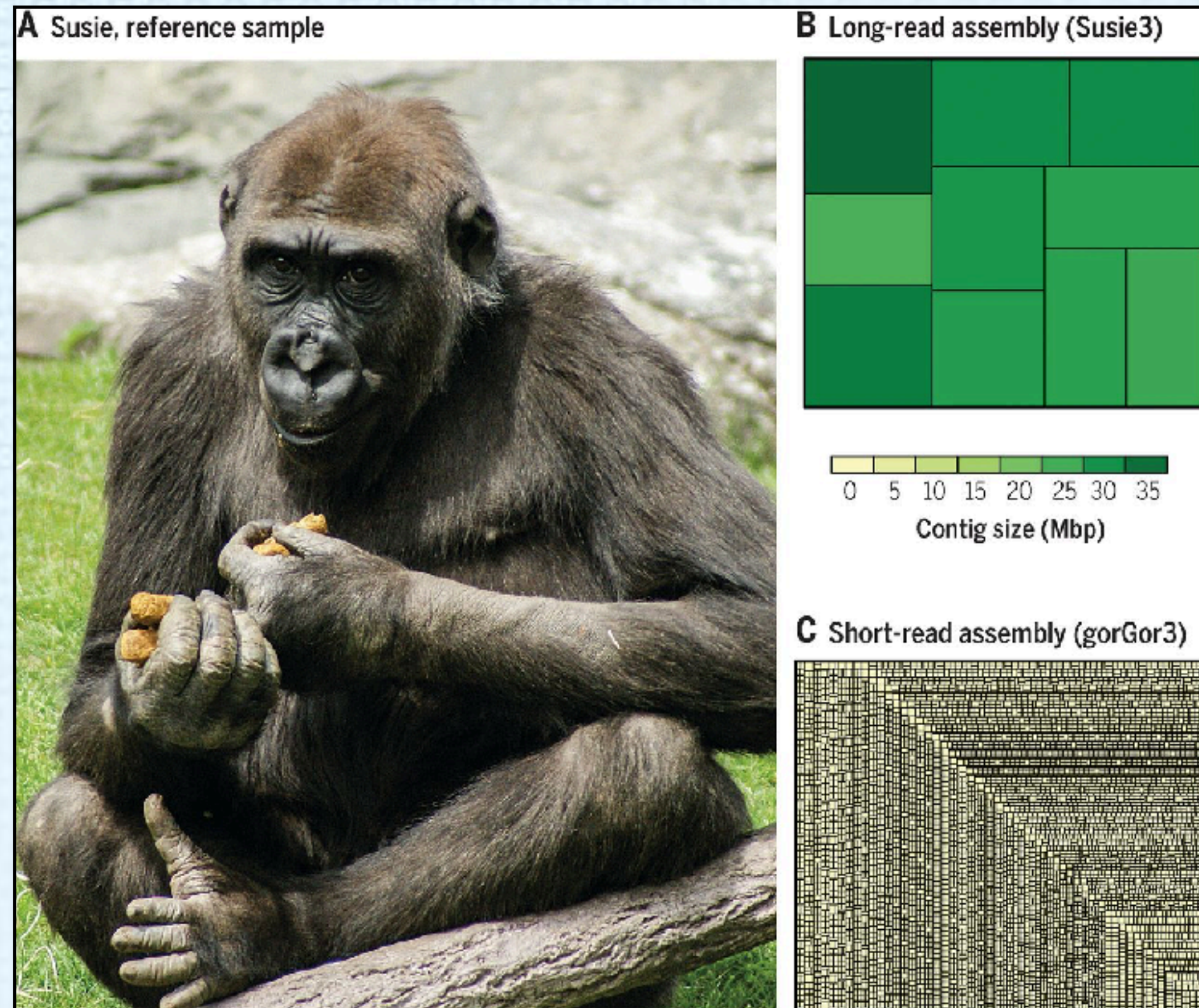
info@phasegenomics.com

Lieberman-Aiden, *et. al.* Science, 2009

Genome assembly does not yield chromosomes

contigs

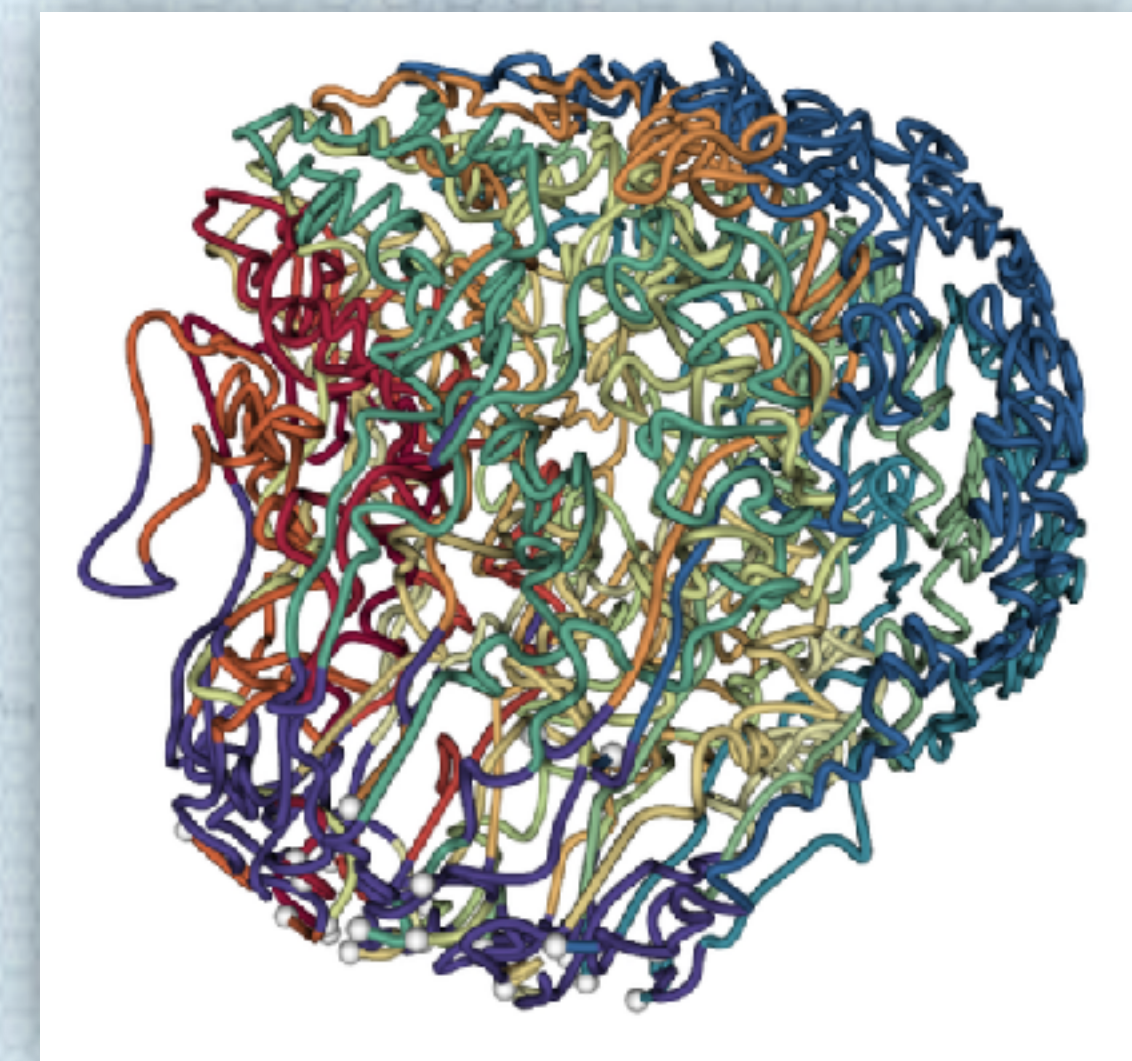
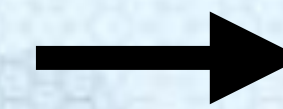
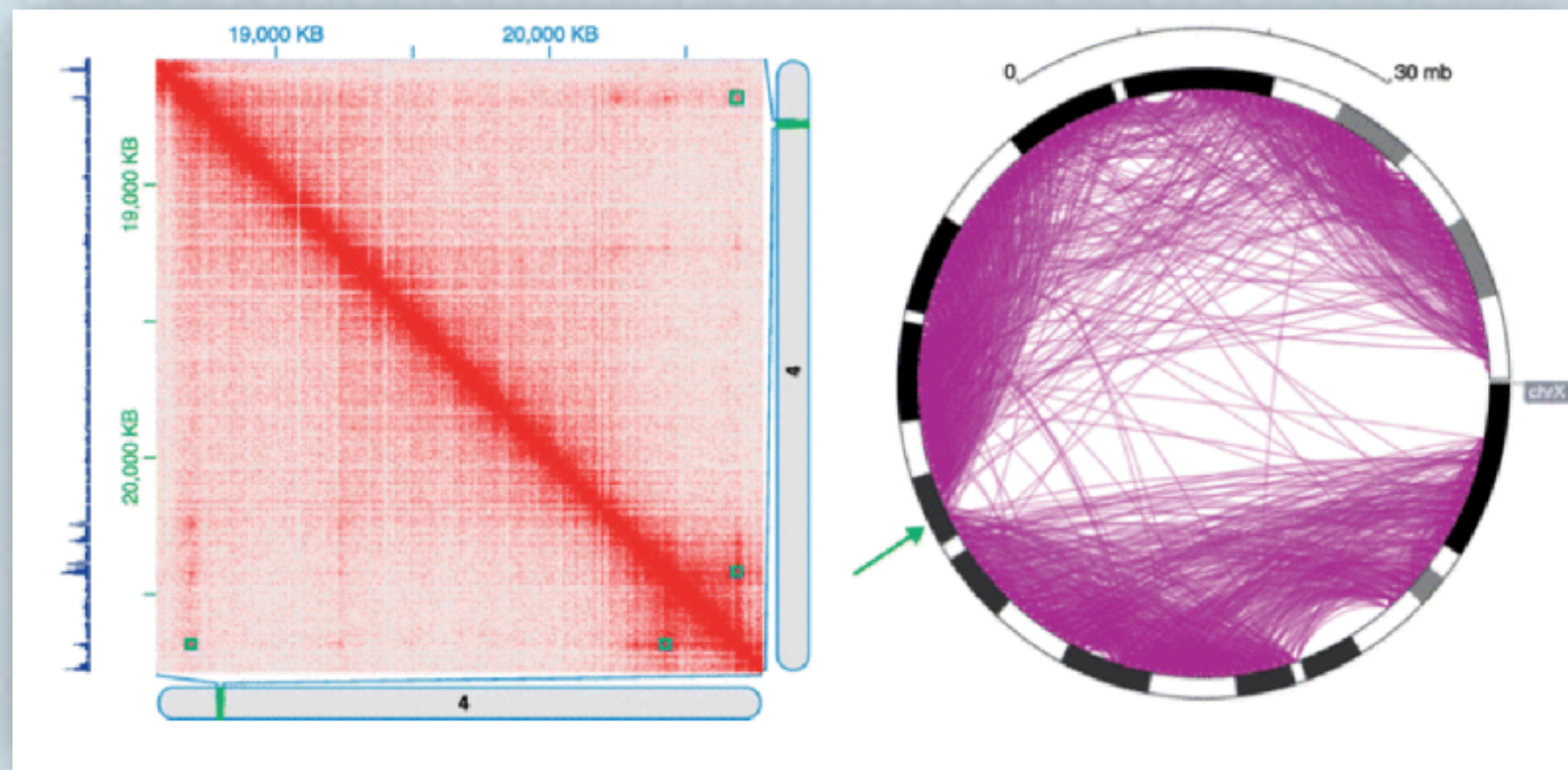
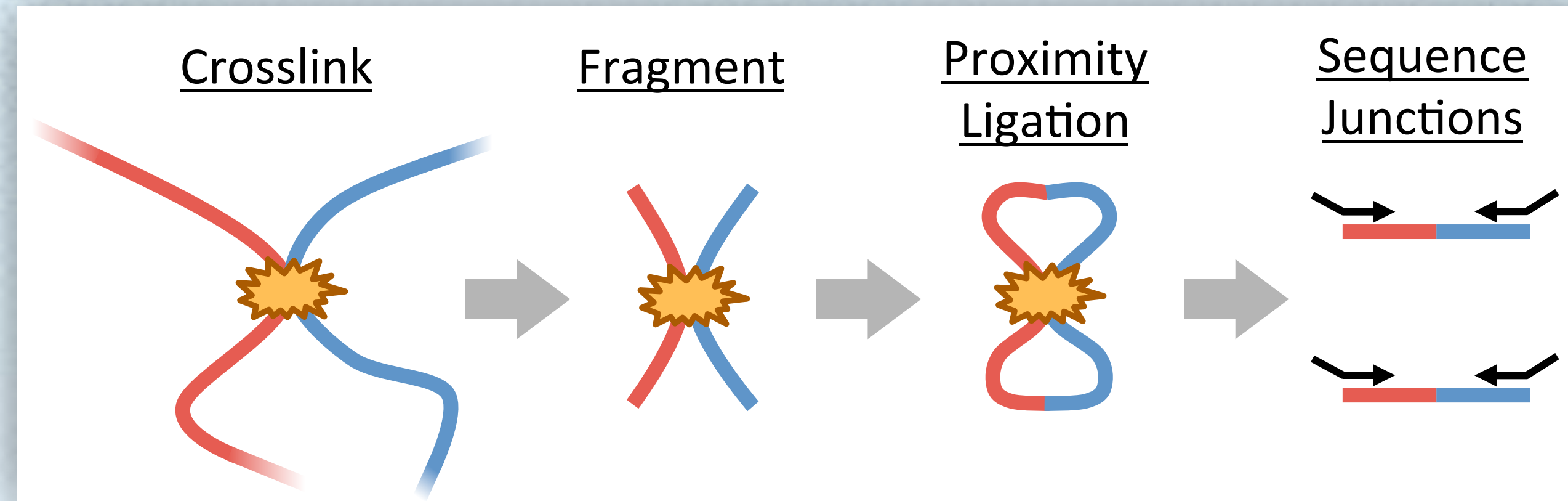
chromosomes



Gordon et al. Science 2016

How do we go from contigs from chromosomes?

Hi-C captures the 3D structure of Chromosomes

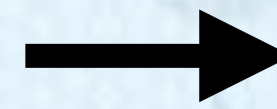


High quality scaffolding with Proximo

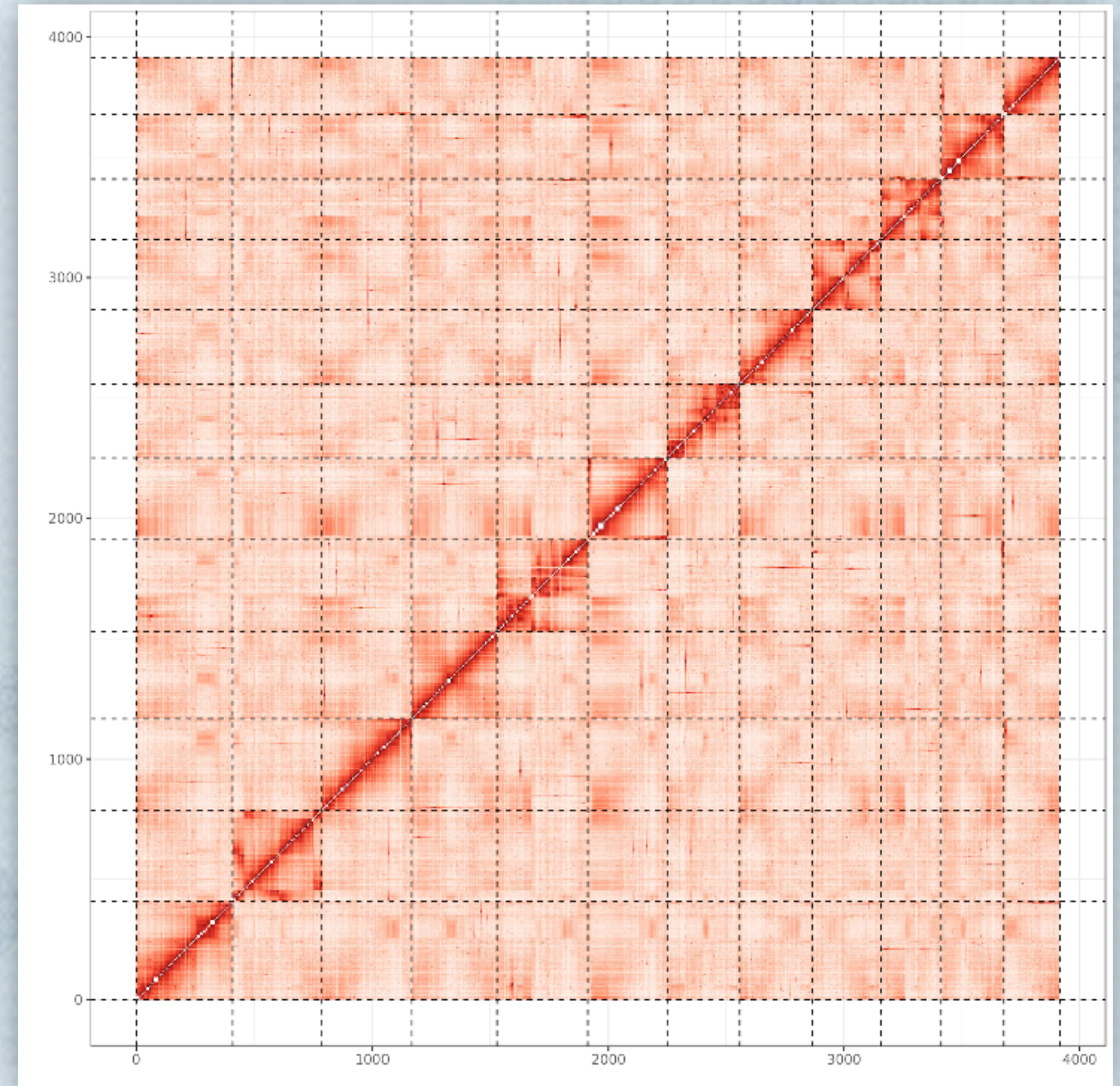
falcon contig to contig proximity



Proximo



chromosome scale scaffolds



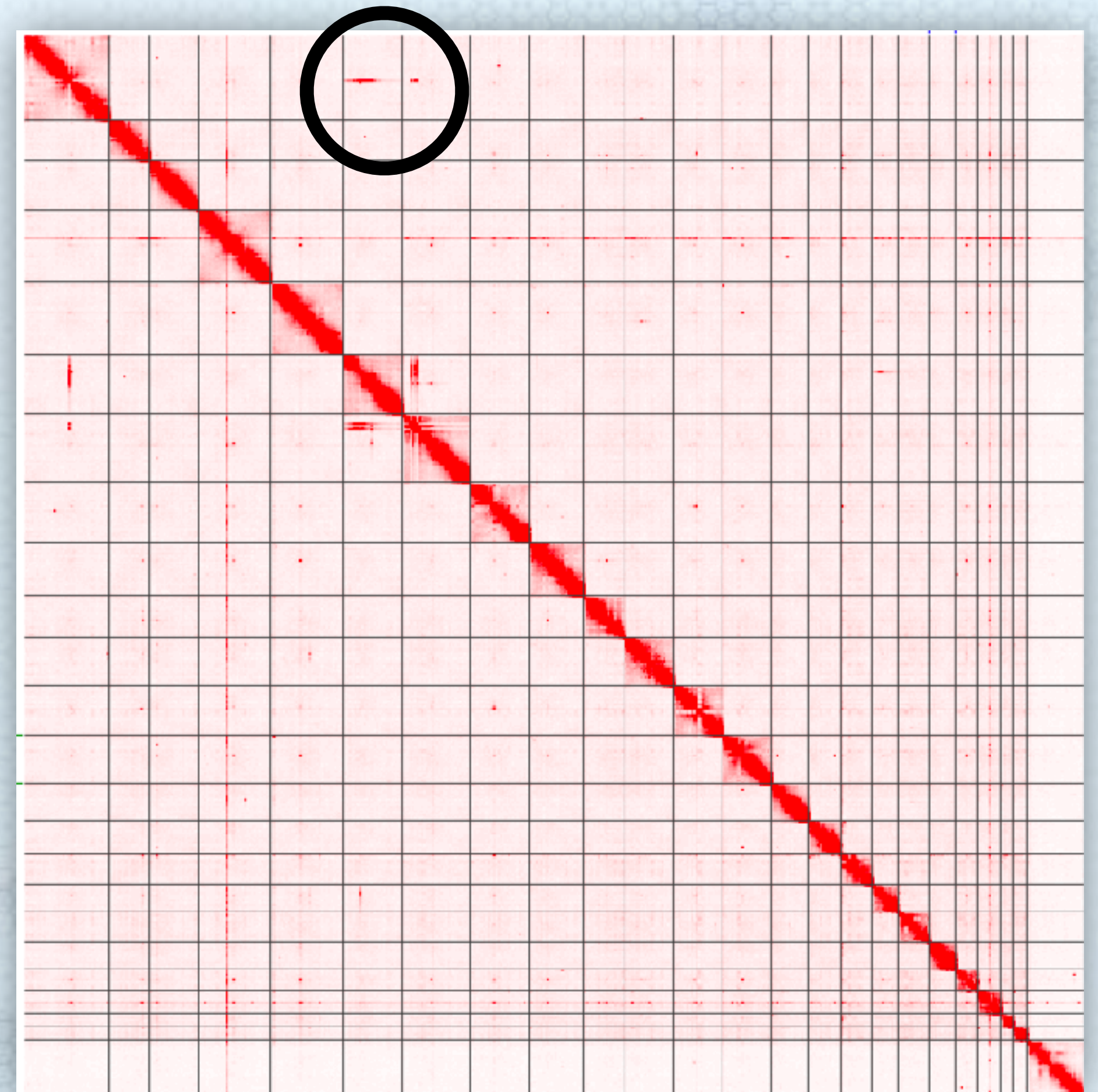
In collaboration with Kelly Vining

Flagging scaffolding issues with Hi-C

Clint the chimpanzee

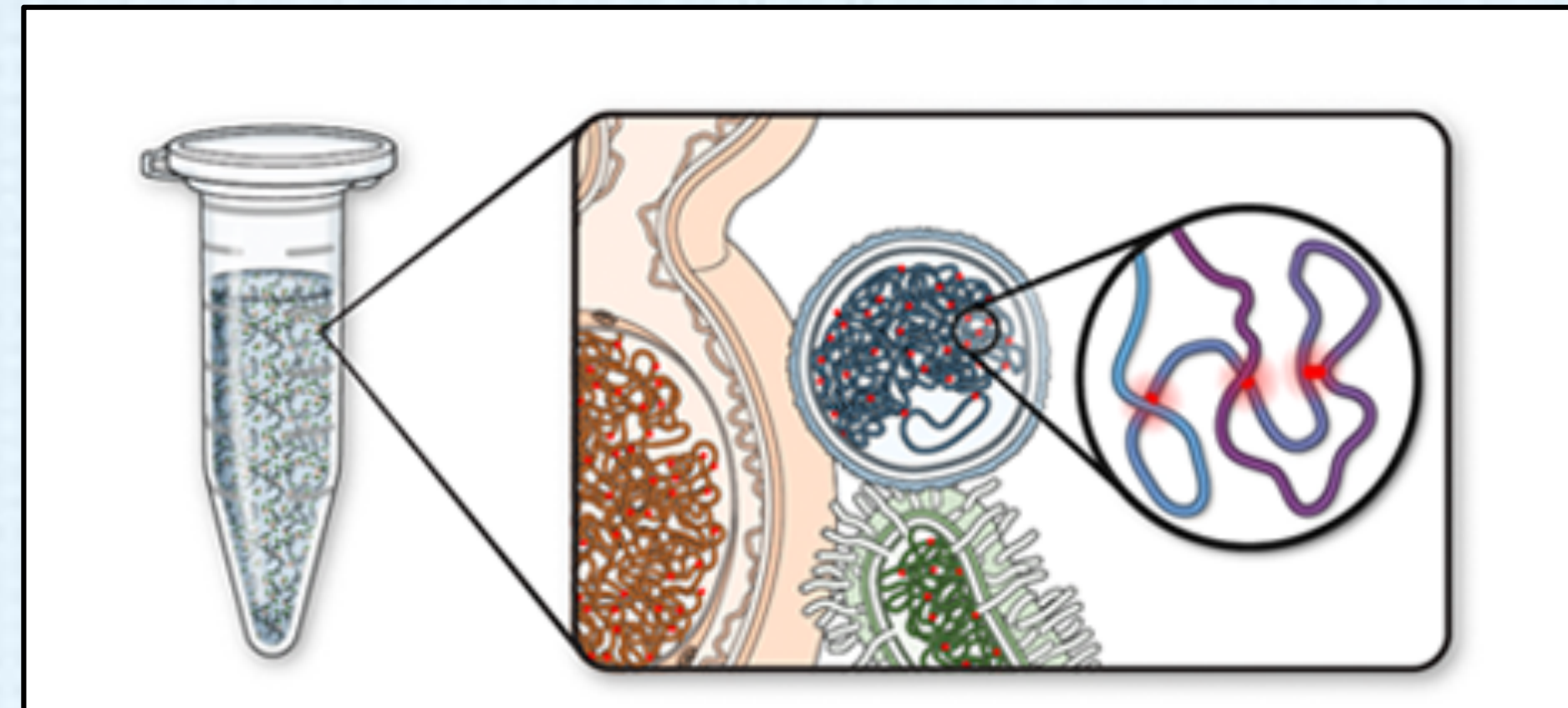


Clint scaffolds



Kronenberg et al., in review

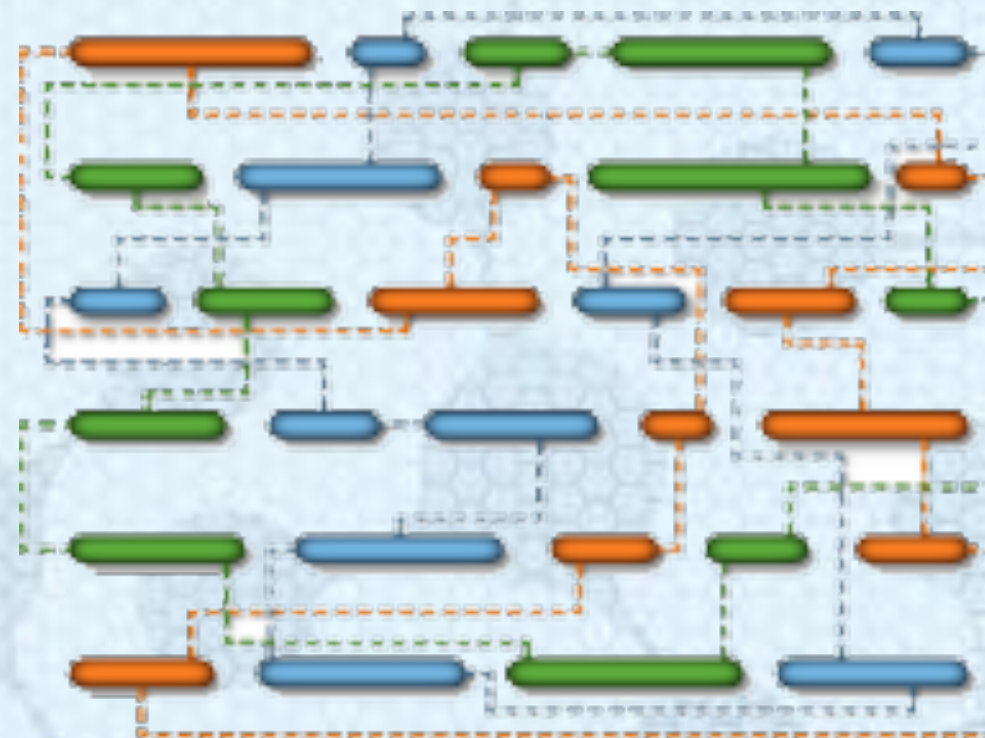
ProxiMeta leverages Hi-C information to deconvolve metagenomic samples



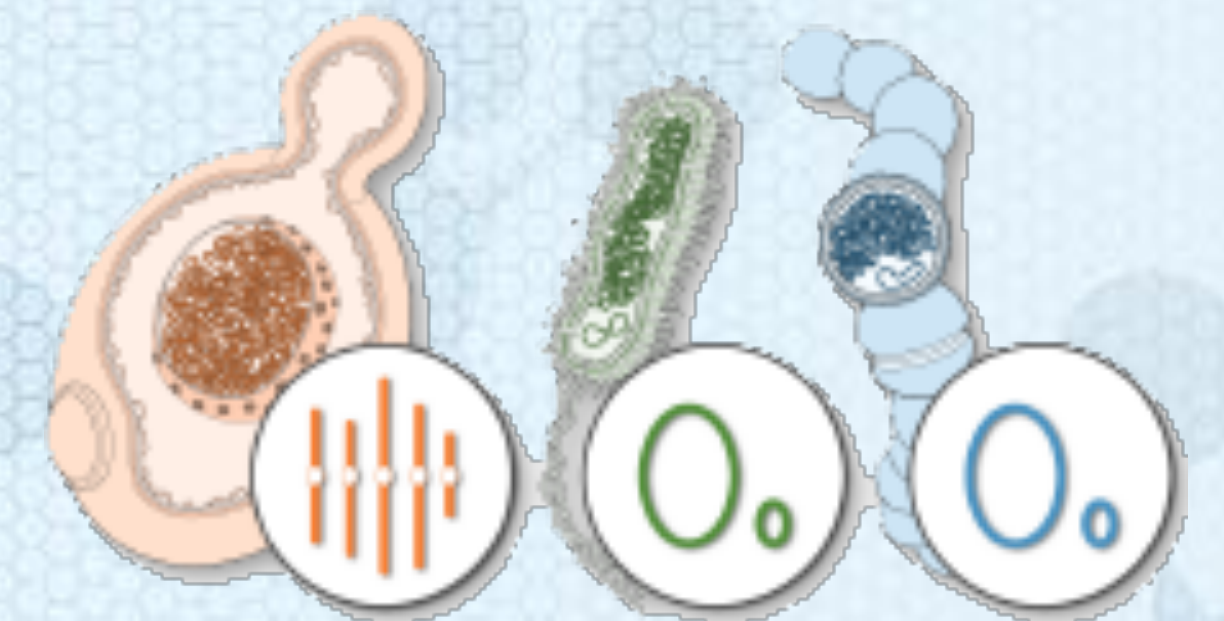
Assembly



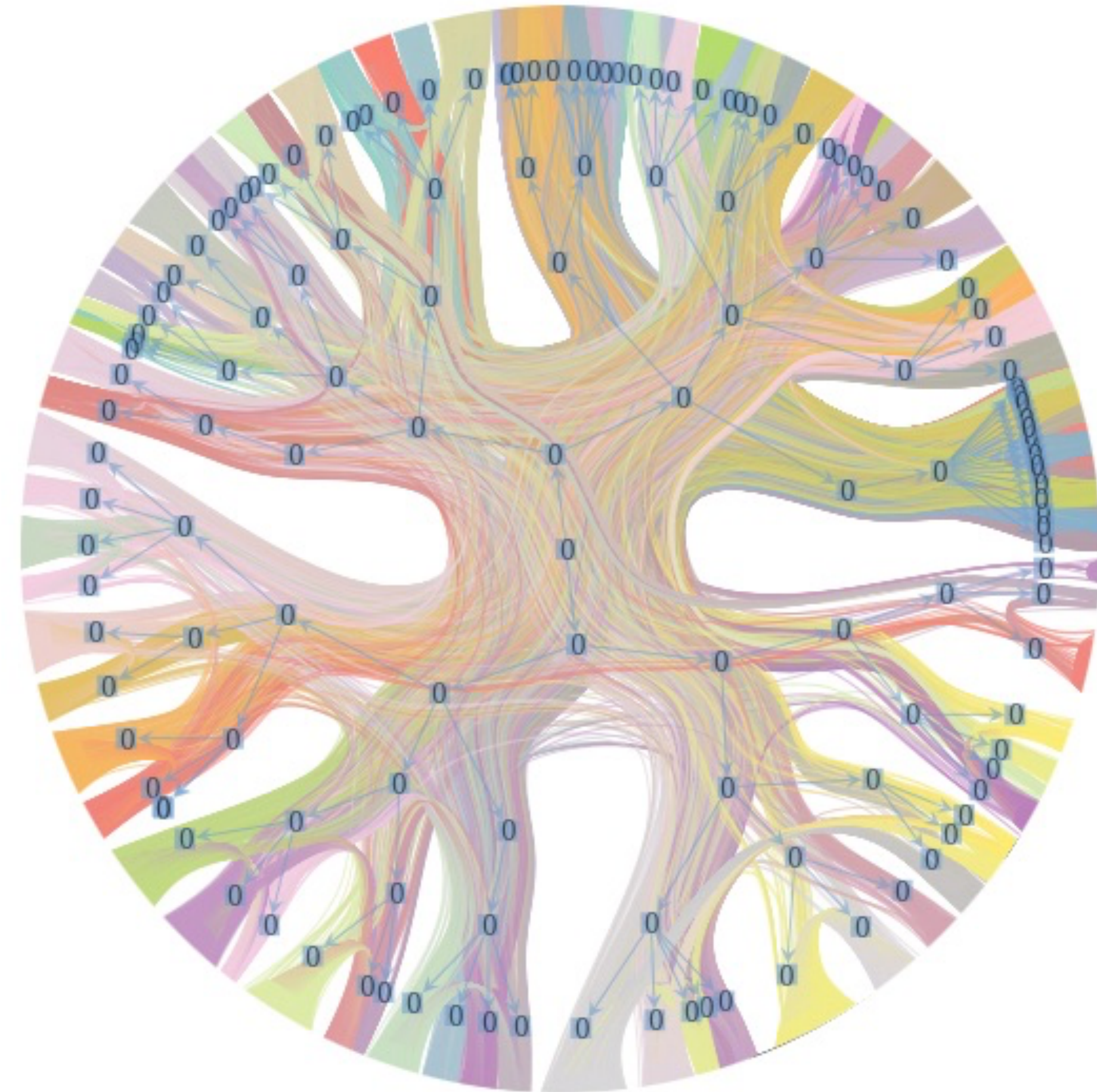
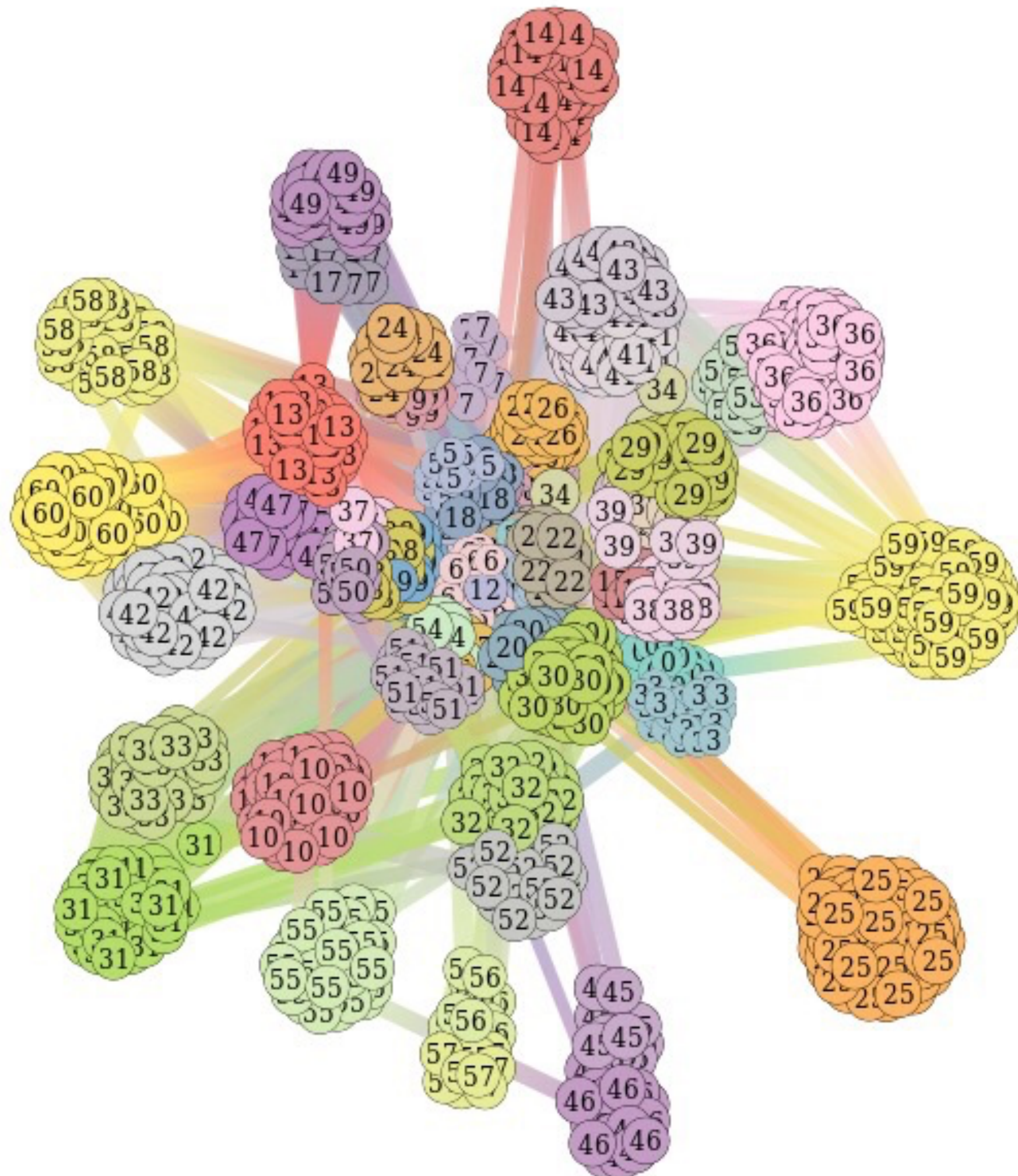
Assembly + Hi-C



Genome clusters



ProxiMeta is a MCMC clustering algorithm to deconvolve metagenomic assemblies



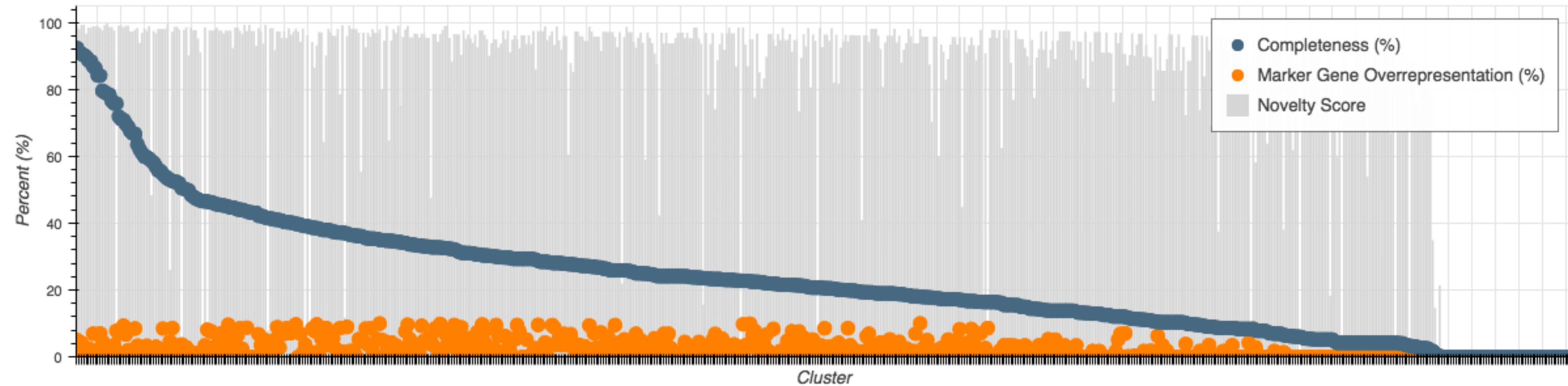
Many genome clusters from complex communities



ProxiMeta Results

Rumen polished pacbio

01/17/2018



Novel Genome
>80% Complete, <10% MGO*
>90 Novelty Score

Known Genome
>80% Complete, <10% MGO*
<90 Novelty Score

Mixed Genome
>80% Complete, >10% MGO*
Any Novelty Score

*Marker Gene Overrepresentation

Cluster ID	Top Reference	Completeness (%)	Marker Gene Overrepresentation (%)	Novelty Score	Abundance (%)	GC (%)	Genome Size	Num Contigs	Contig N50
cluster.1	<i>o__Bacteroidales</i>	92.43	5.08	98.71	-	44.89	2,759,043	13	372,254
cluster.2	<i>o__Bacteroidales</i>	90.93	1.60	97.82	-	48.54	2,484,407	36	77,631
cluster.3	<i>o__Selenomonadales</i>	90.63	4.14	99.12	-	50.26	2,322,633	56	78,971
cluster.4	<i>o__Bacteroidales</i>	90.23	1.80	99.12	-	43.07	2,759,580	12	292,113
cluster.5	<i>p__Euryarchaeota</i>	89.73	1.60	96.48	-	34.86	1,807,690	19	104,356

Collaboration with USDA, NHGRI, Roslin Inst.

kittybiome

luvs

Lil
BUB



Collaboration with AnimalBiome

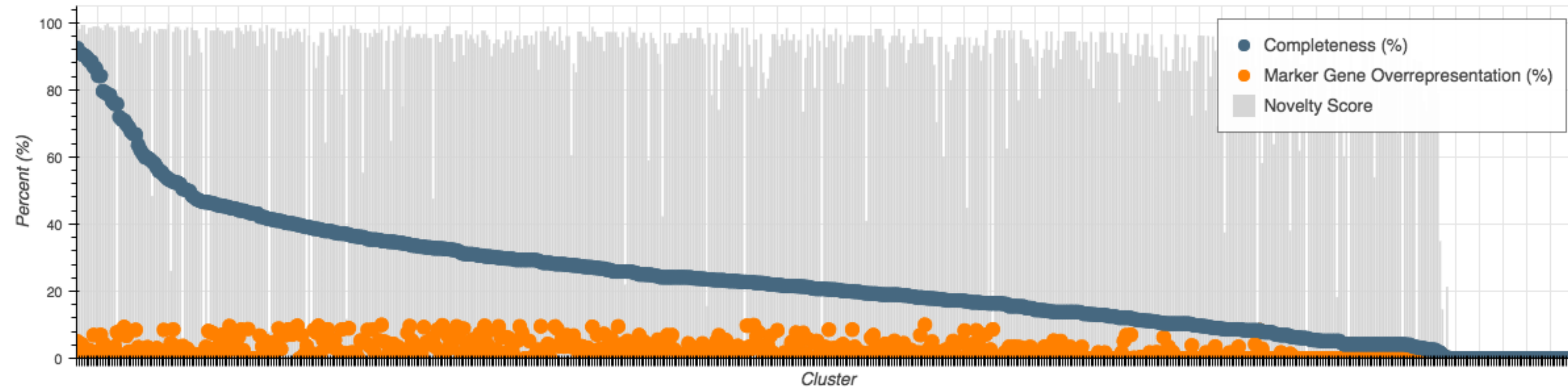
Many genome clusters from complex communities



ProxiMeta Results

Rumen polished pacbio

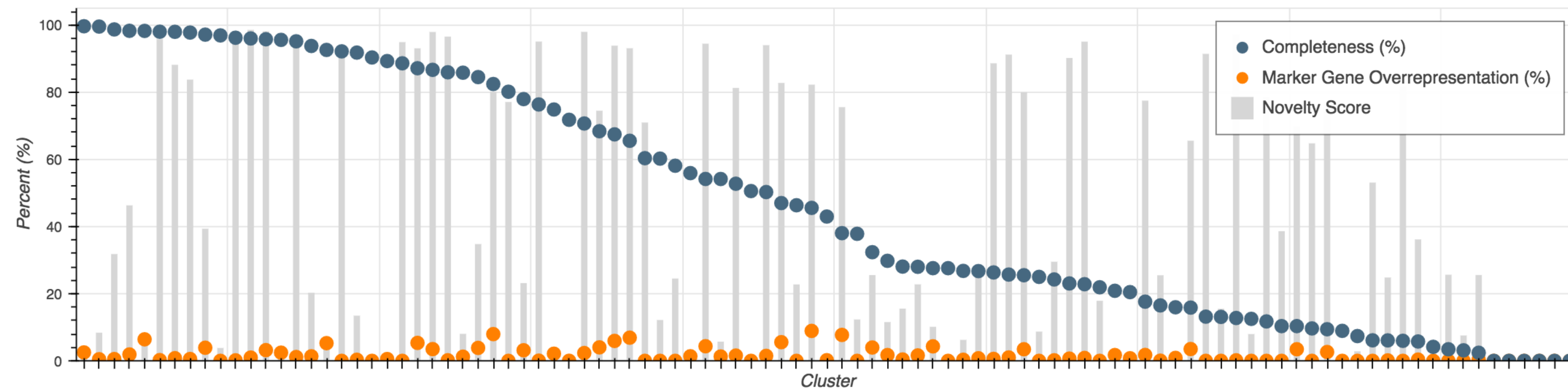
01/17/2018



ProxiMeta Results

Lil bub

01/13/2018



Hi-C low stress



- **Fast (1 day)**
- **Easy to use**
- **Affordable**
- **Simple sample requirements**
- **No HMW DNA**

Metagenomics



Animals



Plants

