My main question is how exactly you would do the analysis. Could you send me an email that spells out all the steps and variables you will measure? I'm looking for details here.

Chapter 3 Outline, Do cell specific effects or chromosome specific effects shape broad scale recombination patterns?

# Background

Wang et al 2019, demonstrates that there is within nucleus covariation for CO count per cell. The proposed mechanism is that within a nucleus, all chromosomes (homologs) have a positive correlation between chromosome axis length, DSBs and COs. This relationship generates an over-dispersed distribution in CO counts per cell, the tail ends of the distribution have cells with more or fewer COs than expected if CO count was purely driven by chromosome effects.

**Main question:**

Is there within nucleus co-variation in CO position / foci spacing (i.e. CO interference)?

Significance**:**

The result of within nucleus covariation for CO number (via SC length) fits with chromatin compaction regulation / DNA loop size and axis length being regulated at a ‘global’ scale within single cells. Whereas, broad scale CO positioning and crossover interference are regulated at a chromosome level (or at least is further removed from global level).

Are broad scale CO positioning patterns driven by single cell effects, like SC length and CO count, or are they driven by chromosome level effects? Answering this question would shed insight on how meiotic processes might be interconnected.

- **Many biological processes are summaries or averages but these processes happen within single cells.**

- A conserved feature of heterochiasmy is the telomere bias in male and uniform CO positing in female yet there are still gaps in understanding this.

Biological significance – for nuc effect on relative location?

-impact?

Different levels (ect)

Among cell variance – how cells vary from each other, Commonly observed over dispersion across many cells. (from chapter1, females have a greater amount)

Within cell variation – (how individual chromosomes differ from each other \*\*Chrm specific effects\*\*

Within cell co-variation – SC length and CO number are positively correlated across bivalents within nuclei

# Testable hypotheses:

# Testable hypotheses:

1. Basic packing ratio predictions: (reflected by within cell covariation of **SC lengths**)
   1. If the packing process (packing ratio) within a cell is uniformly applied across autosomes, then the ratio of var(SC lengths) / var(physical chromosome sizes) should equal ~1. (physical chromosome sizes might not be accurate due to lack of mapping of centromere and telomere sequence which can be an estimated ~10% )
   2. If the packing ratio process does occur uniformly across bivalents within a cell, then the distributions of all pairwise covariance for chromosomes like, cov(Chrm1, Chrm19) ect, will not be significantly different from each other. \*(chromosome effects (deviation from the expected cell-wide packing) could be detected if covariance for a specific chrm were more different than the other pairs.

Packing ratio hypotheses

1. If within cell variance in SC length reflects ‘packing ratio’ (and acts in a uniform/constant cell wide manner), the ratio of var( physical chrm size) : var(within cell SC lengths) should be close to 1.
2. There will be more variance in packing ratio (var(SC lengths = packing ratio) ) than a null expectation (physical size variance) – because all the other metrics have more variance
3. Predictions across metrics (SC lengths, CO counts, CO positions)
   1. The measures of SC length and CO number metrics should have larger **cell effect**, and **smaller chromosome effect (intrinsic)**. CO position and interference should have a smaller nucleus effect and greater **chromosome specific effect**.
4. (differences across mouse groups)

How do the C/A (covariation : total variance) ratios vary between male and female observations?

<I don’t think this female – have greater variance is a good hypothesis to test – the wang paper views it as C/A ratio (covariation : total variance)

* 1. If within cell covariation (is the major component (determinant) of total variation (whole sample among cells / over dispersion – AND can evolve / adjust
  2. ), then in females (which have greater total variance), should also have greater within cell covariance

If the

* 1. the among cell variance for CO number is greater in females, is the within cell covariation is greater? (**sex difference** for within cell covariation)

Over dispersion in CO count commonly observed, females seem to be more over-dispersed compared to males.

1. If within cell covariation in SC / CO number (i.e. “looping factor”) causes greater total variation (within animal) overdispersion, (and this is adaptive) ….

### NOTES

Largest chrm might have the most variance (chrm effects).\*\*how much do I believe the rank.id for my Mus data?

If SC and CO have evolved (in some Musc males)

The within nuc-covariation leads to cell-to-cell variance, (for sc and CO number)

**Wang paper finds that co-variation within cells (looping factor) – contributes a large proportion to the total variation in CO number(total CO per cell) (table 1, C/A ratios ~.5 to .7 for CO number, >.75 for axis length).**

-I’ve been finding mistakes? Or mis-id chrm measures in the provided Wang data

-- also from the methods section, there might be a chance that some of the whole cell (complete FISH) human sperm data is combined with partial FISH biv data

-comparing the ratio of single cell SC variance : physical size chrm variance

I was getting hung up on if chr

There might be 2 versions of covariation

i. from the paper all / the sum all pairwise combinations of chromosomes with covariance calq’d –weight of SC compaction on variance…

ii. single cell covariance (calq covariance all chrms within a cell) – but need two groups for covariance (sum of all pairwise combos --- I don’t think this can be calq’d at single cell level (don’t think this would be meaningful for CO counts.. but maybe sc length. I would call single cell covariance, looping factor / condensation factor

Single cell level of SC length covariation can also be called, ‘looping factor’ (evokes axis-DNA loop organization). The looping factor varies across cells since DNA content remains constant. I don’t think the single cell version can be calculated with co-variance since you need two groups (maybe even-odd type splitting) – but maybe it can be described as with function of ranked SC lengths (focus on SC length) (slope and intercept, these are all close to linear).

I think the looping factors are an important single cell trait / feature – that lead to the over-dispersion / covariance described in the paper. **Wang paper finds that co-variation within cells (combined looping factors) – contributes a large proportion to the total variation in CO number (total CO per cell) (table 1, C/A ratios ~.5 to .7 for CO number, >.75 for axis length).**

Single cell expression and other stuff

-variation in single cell expression levels seems to be adaptive (whittkopp).

Single cell measures: is it adaptable to have (build) in variaiton for a trait across single cells? What are the biological / fitness consequences across the range?

-cells genetically identical

# Approach:

Wang et al 2019 notes

Covariation

(related to my idea of distinguishing between chromosome effect and cell/nuc effect)

**Decomposition of CO variation**

The Wang paper has a big paragraph about decomposing the CO variation (total variation) into 2 main parts; i)CO covariation within a prophase nuc AND ii) intrinsic noise (from sampling all COs into gametes)

Total variance, variance if independence, **covariance**, contribution of covariance to total variance

Overdispersion in the distribution of COs per nucleus directly implies

that the numbers of COs on different chromosomes tend to covary, at higher or lower levels, within individual nuclei (Equation

1 inTable 1). This tendency for covariation canbe observed experimentally.

If chromosomes of an experimental sample are divided

into two comparable groups, e.g., odd-numbered and evennumbered

autosomes (STAR Methods), the numbers of events

on bivalents in the two groups are seen to be correlated within individual

nuclei (Figure 2A). Per-nucleus correlation is also seen

between (1) the numbers of events on two different individual

chromosomes (Figure 2B) and (2) the number of events on one

bivalent with thenumber on all other bivalents in the same nucleus

(Figure S2B). Accordingly, groups of nuclei exhibiting high, intermediate,

or low total numbers of COs exhibit the same hierarchy

ofCOlevels for every chromosome(Figure 2C). Further, forhuman

male, the CV for COs per nucleus and the degree of covariation

between odd- and even-numbered chromosomes exhibit exactly

the predicted relationship (Figure 2E; STAR Methods).

Covariance-

The total variance of CO number per nucleus (‘‘A’’) can be

decomposed mathematically into two sources (Table 1A): the

sum of the ‘‘intrinsic variance’’ of CO numbers on each chromosome

(‘‘B’’), and **the sum of the covariances of CO numbers**

**across all pairs of distinct chromosomes (‘‘C’’).** Values of A, B,

and C can be calculated directly from the experimental data.

The proportional contribution of CO covariance to the total

variance in CO number is given by the ratio C/A. All of the per-nucleus

CO datasets described above exhibit positive values for

total covariance (column C in Table 1B-I). Moreover, the contribution

**of covariation comprises 48%–83% of the total observed**

**variance** (C/A = 0.48–0.83; Table 1B-I) and thus is a major determining

factor in all cases. (Note: the observed covariance cannot

be due to the fact that every pair of homologs almost always

acquires at least one CO [as required to ensure their regular

segregation; Introduction] because, mathematically, neither

variance nor covariance change when one CO is removed from

every bivalent.)

**fig2, E: Cv of total COs per nuc vs per-nuc CO covariation level**

(E) Simulations (STARMethods) confirmthat stronger per-nucleus covariation ofCOs gives increased variation in the total number ofCOs per nucleus (defined byCV).

Sample sizes as in Figure 1. Error bars = SE (A, right, and B). Data sources and details of statistical analysis are given in STAR Methods.

See also Figures S1 and S2.

Tests of within cell summary statistics between empirical data and simulated (randomized data). (much like Wang et al 2019)

Tests / pvalues – permute (Variability index, difference in metric between real and randomized data)

Cell CO count, cell wide total SC

CO position metrics:

-rbar

-average IFD

-average telomere distance

Within cell covariation

Simulate null distribution of relative CO position, keeping the empirical observation of CO counts per cell but draw randomized relative CO positions (from a pool of all empirical positions for a mouse).

**Ideal data set:**

Within all covariation requires whole cell bivalent measures (all 20 chromosomes). Within between cell variance requires a decent sample of cells from each mouse, ~25 good cells. The total cell sample size will limit the sex and strains that are looked at. Generating the whole cell data will require adding in some hand measures for the bivalents the algorithm missed (usually 1-3 missed per cell). This is feasible for a medium dataset, hand measuring 3\*25 bivalents takes about 3 hours. Focus on PWD,MSM, KAZ and a dom strain). What is the effect of rapid evolution of the mean on the between cell variation?

## Results

CO counts and SC lengths have significant cell effects, resulting in real data being over dispersed compared to a randomized dataset. In my preliminary comparisons, the means are the same but the variance changes, for both female and male datasets.

Whereas my results for CO position metrics, when the number of COs and the length of chromosomes for each nucleus is kept constant, but the positions of foci on each chromosome is randomized, the means are significantly different between the real and randomized data.

Random data have more ‘equal spacing’ of foci along chromosomes

--I haven’t tested of the rbar values co-vary within nuclei

**Interpretations**

1. There is within cell covariation in CO position (all data);

-Cell specific effect generate distinct CO position patterns

-SC length regulation and CO position are linked processes (Rapid Prophase Movements? CO:NCO decision)

2. Within cell covariation, but it’s variable across strains or sexes

-Different modes of selection (relaxed vs purifying) via sex specific constraints for CO positioning (spindle or checkpoints)

-Broad scale CO positioning can evolve

3. No covariation for CO positioning; (null = observed)

- CO position is conserved process across the dataset (biophysical constraint for CO:NCO decision)

- Major determinate of CO positing patterns is chromosome specific

**Chapter2 outline notes**

Bivalent structure

-rate of sister cohesion depletion (synchronize)

-something about centromere spring (work done in mitosis yeast?)

Stephens, Andrew D., et al. "Cohesin, condensin, and the intramolecular centromere loop together generate the mitotic chromatin spring." *The Journal of cell biology* 193.7 (2011): 1167-1180.

Thesis Intro – themes

-meiosis as a program, (input-output, bugs, sequential STEPS, checkpoints) useful analogy

(built from mitosis) (some programs run in background, only notice (investigate nuances) when they break